جامعة دمشق
كلية الهندسة الميكانيكية والكهربائية
قسم هندسة الالكترونيات والاتصالات

٣D Face Detection and Recognition

الكشف والتعرف الآلي ثلاثي الأبعاد على الوجه

للمهندس براء الدندن

للمهندس: براء الدندن

إشراف

الدكتور اسماعيل جنود

لجنة الحكم

الأستاذ الدكتور عبد الرزاق البدوية (رئيس القسم)

الأستاذ الدكتور حسن أبو النور

الأستاذ المساعد الدكتور اسماعيل جنود

الفهرس

تمهيد ٣

الفصل الأول دراسة مرجعية لبعض الطرق المستخدمة في التعرف على الوجوه.. ٦

مقدمة ٧

مهام التعرف على الوجوه ٨

خوارزمية التعرف على الوجه ثنائية الأبعاد..... ١٢

نماذج الوجوه ثلاثية الأبعاد ١٧

طرق التعرف على الوجوه ثلاثية الأبعاد ٢٠

قاعدة البيانات المستخدمة ٣٤

الفصل الثاني الكشف وتصحيح الوضعية الآلي للوجوه ثلاثية الأبعاد ٤٣

مقدمة ٤٤

مرحلة المعالجة الأولية(Preprocessing) للوجه ثلاثي الأبعاد... ٤٦

الفصل الثالث التعرف ثلاثي الأبعاد على الوجه ٨٠

التقنية المستخدمة للمقارنة والتعرف ٨١

تسريع عملية التعرف باستخدام تكمية للتمثيل ثلاثي الأبعاد للوجه... ٨٤

تجزيء الوجه والتعرف عليه ٨٦

المقارنة(فحص التطابق) ٩١

الفصل الرابع نتائج عملية التعرف ١٠٠

قائمة بالمفردات المستخدمة وترجمتها ١٢٤

قائمة بالاختصارات المستخدمة وترجمتها ١٣٠

المراجع ١٣١

تمـهيـد

التعرف الآلي على وجه الانسان له تطبيقات عديدة. وفي حين تم انجاز الكثير من التقدم في التعرف ثنائي الأبعاد على الوجه مع وجود الألوان فإن التقدم في تقنية التعرف ثلاثي الأبعاد على الوجه أقل كثيراً.

إن التعرف الدقيق والقوي على الانسان مهم في كثير من المسائل. تم البحث خلال السنوات السابقة في عدة طرق علمية للتعرف على الأشخاص. هناك طريقة مدروسة على نطاق واسع وهي biometrics الخواص البنيوية أوالفيزيولوجية أوالسلوكية التي يعتقد أنها وحيدة ومتميزة لكل شخص. فيليبس لخص خواص نظام التعرف المثالي المعتمد على الbiometrics :

– كل أعضاء المجموعة يجب أن يكون لها الbiometric المستخدم في عملية التعرف.

–البصمة البيومترية(الخواص البيومترية) لكل شخص يجب أن تختلف عن بقية الأشخاص في المجموعة المدروسة

– الخواص البيومترية لكل شخص يجب أن لا تتغير تحت الظروف التي يتم فيها الحصول على هذه الخواص.

_ يجب على النظام أن يقاوم الظروف التي قد تخفض من أدائه.

الطريقة العلمية الأولى للتعرف على الأشخاص بالاعتماد على القياسات البنيوية للإنسان كانت نظام Bertillon الذي طرح في فرنسا في القرن التاسع عشر. في هذا النظام تم استخدام القياسات البنيوية مثل عرض الجمجمة و طول القدم ولون الشعر ولون العين والصور الأمامية والجانبية.

تم التوقف بالعمل بهذه الطريقة بسبب عدم قدرتها على القيام بتعرف دقيق وتم استبدالها بأنظمة أخرى أكثر وثوقية تعتمد على بصمات الأصابع. التحقق باستخدام بصمات الأصابع كان يعتبر الطريقة الأكثر وثوقية للتعرف على الأشخاص خلال قسم كبير من القرن العشرين وتم استبدالها فيما بعد بطريقة التعرف والتحقق باستخدام DNA .

مع توفر الحواسب كان من الطبيعي التطوير للوصول إلى التعرف الآلي على الناس. في السنوات الأخيرة وبسبب توفر الحواسب وزيادة سرعتها وحجم ذاكرتها أصبحت أنظمة التعرف البيومترية الآلية مجالاً هاماً للبحث ولهذه الأنظمة تطبيقات عديدة منها الدخول المؤمن إلى آلات ATM وإلى المباني ومن هذه الاستخدامات أيضاً :

المراقبة الآلية – التحليل الجنائي – معالجة الصور من قواعد البيانات لدى أقسام الشرطة – التعرف الآلي على المرضى في المستشفيات – مراقبة عمليات السرقة – التفاعل بين الانسان والحاسب .

الأنظمة البيومترية الآلية تسرع عملية معالجة الحجوم الكبيرة من البيانات وفي بعض الحالات قد تعمل في الزمن الحقيقي. حالياً ومن أجل التعرف تستخدم بشكل شائع التقنيات التالية :الأرقام المعرفة الشخصية – بطاقات\كودات الولوج – tags تعريف عن طريق الترددات الراديوية. أرقام التعريف وبطاقات\كودات الولوج تحتاج أيضاً تدخلاً بشرياً كبيراً وبالتالي فهي ذات استخدام قليل في التعرف على الأطفال الصغار والأشخاص المرضى.

عدة تقنيات بيومترية تم البحث فيها من أجل التعرف الآلي على الانسان منها الوجه والعين وبصمات الأصابع وبصمة راحة الكف واليد والصوت وطريقة المشي والتعرف على الخط أثناء الكتابة.

من بين هذه الأساليب تبين أن أنظمة التعرف باستخدام بصمات الأصابع و باستخدام العين دقيقة جداً لكنها تتطلب تعاوناً(تدخلاً) كبيراً من قبل المستخدم وهي صعبة التطبيق في المراقبة في الزمن الحقيقي حيث يكون من المرغوب أن يكون التدخل البشري أقل ما يمكن أو يكون من المرغوب القيام بالمراقبة السرية. التعرف على الوجوه يحتاج أقل تدخل من قبل الانسان وملائم من أجل تطبيقات المراقبة ويمكن تنفيذه باستخدام مكونات رخيصة نسبياً.

على الرغم من المهارة العالية لدى الانسان في التعرف على الوجوه فإن هناك أيضاً بعض الاخفاقات في قدرات الناس في التعرف على الوجوه. أبحاث الDNA أظهرت أن ٨٤% من حالات الادانة الخاطئة كانت جزئياً بسبب التعرف الخاطئ من قبل الشهود أو الضحايا. يعتقد الباحثون ان قدرات التعرف على الوجوه لدى الانسان متاثرة ببعض العواطف. أكثر من ذلك كثير من الخواص الفيزيولوجية لمعالجة وجوه الناس والتعرف عليها غير مدروسة بشكل جيد.

كمثال آخر لدينا مجال مكافحة الجريمة حيث يتم عرض ست صور وجوه أمام الشهود أو الضحايا ليتعرفوا عليها وأحد هذه الصور هي صورة المشتبه به. من المفضل بالاضافة إلى صورة وجه المشتبه به أن تكون باقي الصور لأشخاص قريبين في المنظر للمشتبه به .

من أجل المقارنة الفعالة من الضروري تعريف قياسات فعالة لأوجه الشبه بين الوجوه. يمكن لمثل هذه القياسات أن تكون مفيدة أيضاً في حالة معالجة الصور من قواعد بيانات الوجوه. يؤكد العلماء الحاجة لتعريف قياسات كمية لأوجه الشبه بين وجوه الناس ويعتقدون أن استخدامات الحاسب في معالجة الصور تساعد في فهم الآليات التي تحكم عمليات معالجة صور الوجوه .

يحاول هذا البحث تسليط الضوء على المساهمة التي يمكن أن تقدمها تقنية التعرف ثلاثية الأبعاد في مجال التعرف على الوجوه فقد تضمن الفصل الأول مسألة التعرف على الأشخاص وطرق التعرف ثنائية الأبعاد المطروحة وتم توضيح المشاكل التي تعاني منها هذه الطرق وكذلك التقنيات المستخدمة للحصول على نماذج وجوه ثلاثية الأبعاد في الزمن الحقيقي وكذلك بنية البيانات المستخدمة لتمثيل هذه النماذج وتوضيح بعض طرق التعرف على الوجوه ثلاثية الأبعاد. وكذلك تم شرح خصائص قاعدة البيانات المستخدمة في هذا البحث من صور الوجوه ثلاثية الأبعاد وكيفية الحصول عليها وخصائصها وظروف التقاطها ومشاكل هذه الصور وطرق معالجتها.

أما الفصلين الثاني والثالث فقد تضمنا تصميم كافة الخوارزميات اللازمة لاقتطاع منطقة الوجه المطلوبة وتصحيح وضعيته(pose correction) وايضاح نتائج هاتين الخطوتين ثم القيام بمحاولة التغلب على تأثيرات تغير التعابير في الوجه باقتطاع المناطق الأقل تأثراً بهذه التعابير(واستعراض أسلوبين ممكنين لذلك) وبناء قاعدة معطيات منها لاستخدامها في عملية التعرف ثم توضيح الأسلوب المتبع للتعرف على الوجوه بالوصول إلى الخطأ الأصغري بين الوجوه المقارنة بواسطة تعويض الانزياح والدوران بينها وكذلك تسريع عملية التعرف باستخدام تكمية للتمثيل ثلاثي الأبعاد للوجه.

أما الفصل الرابع فقد تضمن نتائج عملية التعرف من أجل حالات مختلفة(وجوه فيها ابتسامة أو ضحكة ووجوه عادية بدون تعابير ووجوه تتضمن أو لا تتضمن تغيرات في الوضعية أي دوران حول المحاور) وتوضيح الاستنتاجات المتعلقة بهذه النتائج ومقارنتها من أجل أساليب مختلفة لاقتطاع المناطق الأقل حساسية للتعابير.

الفصل الأول

دراسة مرجعية لبعض الطرق

المستخدمة في التعرف على الوجوه

١-١ مقدمة:

هناك الكثير من جهود البحث التي بذلت خلال العقدين الماضيين لتطوير أنظمة أتوماتيكية قوية للتعرف على الوجوه. بشكل عام ركزت جهود البحث بشكل أساسي على الصور ثنائية الأبعاد الملونة وهناك أنظمة تجارية متاحة الآن للاستخدام.

أنظمة التعرف على الوجوه ثنائية الأبعاد(٢D) سهلة البناء وبمكونات رخيصة لكنها لا تؤمن تعرفاً قوياً على الوجوه. اختبار مصنعي أنظمة التعرف على الوجوه تم انشاؤها عام ٢٠٠٢ (FRVT) لوضع بارامترات لأداء خوارزميات التعرف على الوجوه الاتوماتيكية ثنائية الأبعاد. وتم ملاحظة أن أداء أفضل ثلاث خوارزميات انخفض تقريباً إلى النصف من أجل صور الوجوه الملتقطة تحت ظروف اضاءة مختلفة أو بوضعيات مختلفة للرأس. الصور الأمامية ثنائية الأبعاد المركبة باستخدام نماذج قابلة للتشكيل ثلاثية الأبعاد حسنت كثيراً نتائج عملية التعرف من أجل الوجوه ذات التغير الكبير في الوضعية pose. وبالتالي بالنسبة ل FRVT ٢٠٠٢ اعتبر أنه باستخدام نماذج وجوه ثلاثية الأبعاد من أجل تصحيح الوضعية pose correction يمكن حل مشكلة الوضعية في عملية التعرف على الوجوه. المرجع [٢]

تقنية التعرف على الوجوه ثلاثية الأبعاد ظهرت في السنوات الأخيرة وخاصة بسبب توفر أجهزة مطورة لالتقاط الصور ثلاثية الأبعاد. بالنسبة لخوارزميات التعرف على الوجوه ثلاثية الأبعاد يتم استخدام نماذج الوجوه ثلاثية الأبعاد بحد ذاتها (خوارزميات ثلاثية الأبعاد ٣D) أو معاً مع الصور ثنائية الأبعاد (خوارزميات ٢D+٣D).

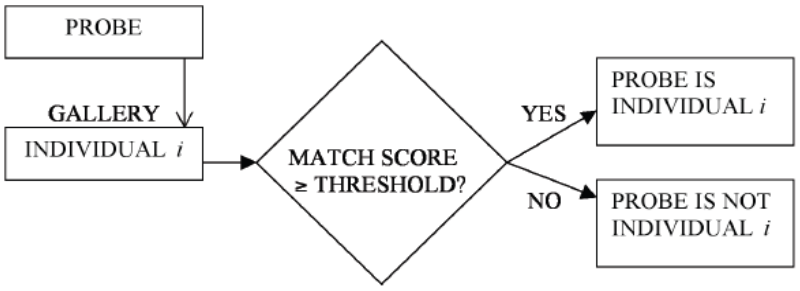
نوضح أولاً المهام الرئيسية التي يتم انجازها بواسطة أنظمة التعرف الأتوماتيكية والقياسات الكمية لقياس أدائها. بعد ذلك سنناقش الخوارزميات الرئيسية ثنائية الأبعاد للتعرف على الوجوه التي أثرت على تطوير تقنيات التعرف على الوجوه ثلاثية الأبعاد وتم استخدامها في كثير من طرق التعرف (٢D+٣D). وثلاثية الأبعاد(٣D) المستخدمة حالياً للتعرف على الوجوه .

١-٢ مهام التعرف على الوجوه :

المهمتان الأساسيتان اللتان يتم انجازهما من قبل أي نظام تعرف بشري هما التحقق (verification\authentication) والتعرف (تحديد الشخصية Identification) وسيتم مناقشتهما.

١-٢-١ التحقق Verification :

التحقق هي عملية مقارنة واحد مع واحد one-to-one حيث يدعي شخص أنه أحد الأشخاص المعروفين للنظام (الشكل ١-١). قاعدة البيانات للأشخاص المعروفين للنظام تسمى "gallery" .

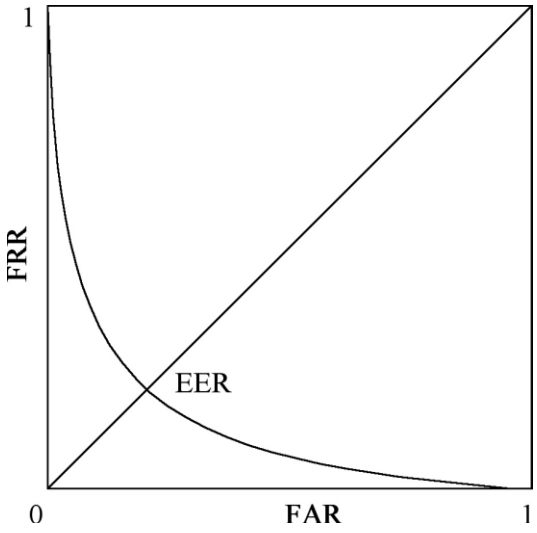


الشكل (١-١) مخطط لنظام تحقق بيومتري اتوماتيكي

الشخص الذي يتم التحقق من شخصيته من قبل النظام يشار إليه "probe" التمثيل الوجهي للprobe يتم مقارنته مع التمثيل لهذه الشخصية في قاعدة البيانات"gallery". إذا كانت محصلة التشابه بين الاثنين أكبر من عتبة معرفة مسبقاً فإنه يتم التعرف على الشخص على أنه فعلاً الشخص الذي يدعي وإلا فإنه يعتبر غير موجود في قاعدة البيانات. كمثال عن ذلك سيناريو تحقق الدخول الآمن إلى بناء معين.

أداء أنظمة التحقق يتم التعبير عنها بواسطة منحني خواص عمل مستقبل receiver operating characteristic curve(ROC). منحني ROC هو رسم يوضح العلاقة بين معدل القبول الخاطئ false acceptance rate(FAR) ومعدل الرفض الخاطئ false rejection rate(FRR). عملية التحقق يمكن النظر إليها على أنها مسألة ذات مستويين حيث محصلة التطابق يمكن أن تصنف على أنها داخل المجال المقبول ليتم التحقق من الشخص على أنه الشخص الذي يدعي أو قد تكون محصلة التطابق خارج هذا المجال.

تعرف FAR على أنها نسبة المقارنات بين شخصين مختلفين والتي تقبل بشكل خاطئ على أنها تطابق. FRR نسبة المقارنات بين صورتين للشخص نفسه والتي يتم بشكل خاطئ رفضها من قبل النظام. كل من FAR و FRR تتغيران مع تغير عتبة القرار للنظام. مقياس أداء وحيد مخصص لأنظمة التحقق هو معدل الخطأ المتساوي equal error rate(EER) حيث FAR=FRR. من أجل نظام مثالي يكون EER=٠٪. المنطقة تحت منحني ROC تعد أيضاً مقياساً للأداء. تتراوح هذه المنطقة من الصفر من أجل نظام مثالي إلى ٠,٥ من أجل نظام غير مثالي. الشكل ١-٢ يمثل منحني ROC نموذجي. نلاحظ أنه تم استخدام منحنيات ROC متكافئة رغم اختلاف أشكالها في تقييم أنظمة التصنيف في تطبيقات أخرى(مثلاً في التصوير الطبي). المرجع [٢]



الشكل(١-٢) منحني خواص عمل مستقبل

منحنيات ROC مرتبطة أيضاً بمنحنيات الاستعادة-الدقة precision-recall التي تستخدم كثيراً في تقييم أداء أنظمة استقبال المعلومات العامة. الاستعادة مساوية لنسبة الموجب الصحيح (true positive) والتي تقيس نسبة النتائج (القيم scores) التي تحقق التطابق(أصغر من العتبة) والتي تم اعتبارها تطابقاً بشكل صحيح من قبل النظام و الدقة من جهة اخرى تقيس نسبة قيم التطابق والتي يتم اعتبارها من قبل النظام كقيم خارج عتبة التطابق وهذا الاعتبار يكون صحيحاً .

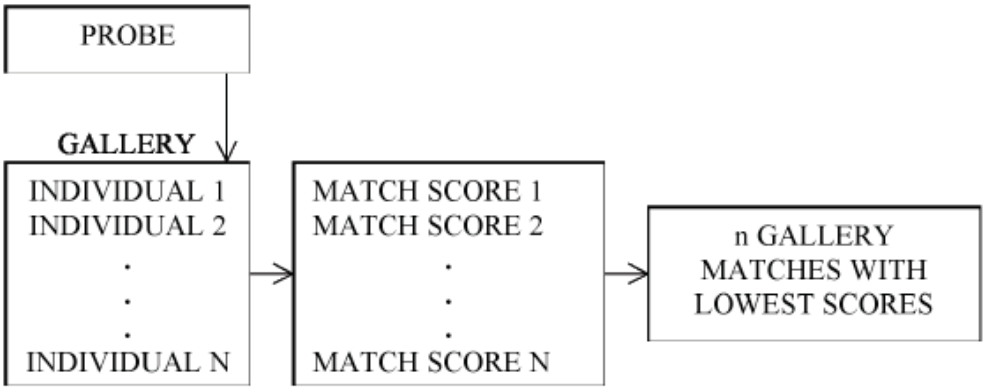
منحنيات الاستعادة – الدقة عموماً أخذت دور منحنيات ROC لتقييم أداء أنظمة القرار ثنائي الصنف مع مجموعات بيانات فيها تغيرات كبيرة وظروف التقاط قاسية إلى حد بعيد. مجموعة البيانات المستخدمة لتقييم أنظمة التحقق من الوجوه أيضاً فيها ظروف التقاط صعبة للصور طالما ان عدد قيم التطابق الأكبر من العتبة عادة أقل بكثير من عدد قيم التطابق الأصغر من العتبة (تطابق مع أحد الشخصيات في قاعدة البيانات). لكن في الوقت الحالي تستخدم فقط منحنيات ROC لتقييم أدائها وهكذا فإن دراسة استخدام منحنيات الاستعادة – الدقة من أجل هذه المهمة يعتبر مبرراً .

١-٢-٢ التعرف Identification :

التعرف عملية بحث عن تطابق بين واحد و أكثر من شخص (one to many) حيث يتم التعرف على شخص غير معروف بمقارنة البيومترك الخاص به أو بها مع قاعدة بيانات للأشخاص المعروفين. يتم ايجاد التطابقات الأقرب في قاعدة البيانات"gallery". مثلاً قد يتم التعرف إلى شخص بمقارنته مع قاعدة بيانات في قسم للشرطة. تعتبر المراقبة لقائمة تعميماً لعملية التعرف. في عملية المراقبة في قائمة يتم مقارنة كل شخص probe مع ملامح كل الأشخاص المعروفين للنظام ونحصل على أعلى n قيم للتشابه التي تكون فوق عتبة معينة وتعتبر تطابقات. يعتبر وضع قيمة عتبة معرفة مسبقاً كحماية كي لا نختار أشخاص من قاعدة البيانات "gallery" من بين أكثر n تطابق دون أن يكونوا قريبين في المنظر من الشخص تحت الاختبار probe. يمكن تقييم أداء نظام تعرف بواسطة منحني خواص محصلة التطابق cumulative match characteristic (CMC) هذا التشكيل يفترض أن النموذج مغلق حيث كل الأشخاص الذين يتم اختبارهم موجودون في قاعدة البيانات "gallery" .

منحني CMC يوضح علاقة معدل التعرف Recognition Rate(RR) مع أعلى n تطابق مع قاعدة البيانات. انها نسبة عدد الأشخاص تحت الاختبار probe الذين يكون التطابق الصحيح في قاعدة البيانات موجوداً ضمن أكثر n تطابق إلى العدد الكلي للأشخاص الذين يتم مقارنتهم مع قاعدة البيانات. وإذا لم يكن صحيحاً افتراض أن النموذج مغلق أي أن أحد الأشخاص تحت الاختبار probe غير موجود في قاعدة البيانات gallery فعندئذ تكون RR العظمى التي يمكن الحصول عليها أقل من ١٠٠% .

التحققVerificationوالتعرفIdentification وعمليات مراقبة القوائم تفرض تحديات مختلفة في التصميم. مثلاً قد يحاول بعض المنتحلين خداع نظام التحقق والتحكم بالدخول مغيرين شكلهم إلى شخص ما معروف لدى النظام وبالتالي من المهم لنظام التحقق والتحكم بالدخول الموثوق والآمن أن يكونFAR له منخفضاً كي لا يسمح بالخطأ بدخول منتحل لشخصية ما معروفة للنظام. بالمقارنة فان الناس الذين يحاولون تجنب عملية التعرف نادراً ما يحتاجون تغيير شكلهم إلى شخص آخر. وهكذا، مثلاً، فإنFRRلأنظمة المسح في المطار يجب أن تكون منخفضة كي تسمح بالتعرف إلى الأشخاص رغم وجود بعض التغير في شكلهم وهكذا فإنه من المهم معرفة نوع التطبيق أثناء تصميم الأنظمةالبيومترية فالنظام المثالي من أجل تطبيق معين قد يكون أداؤه غيرمقبول بالنسبة لتطبيق آخر.



الشكل (١-٣) مخطط صندوقي لنظام تعرف بيومتري آلي

١-٣ خوارزميات التعرف على الوجوه ثنائية الأبعاد :

تقنية التعرف على الوجه ثنائية الأبعاد تطورت بشكل كبير خلال العقدين الماضيين. ونقول بشكل عام أن الطريقتين العامتين للتعرف على الوجوه ثنائي الأبعاد هما :

a: تعتمد على المظهر الكلي للوجه.

b: تعتمد على قوالب لملامح الوجه المحلية.

خلال العقد الماضي تم إجراء عدد من الاختبارات المستقلة لمقارنة أداء خوارزميات التعرف على الوجوه ثنائية الأبعاد. تضم هذه الاختبارات سلسلة اختبارات (تقنية التعرف على الوجوه) FERET و اختبارات (مصنعي التعرف على الوجوه). خلال عمليات التقييم هذه أظهر قليل من خوارزميات التعرف على الوجوه ثنائية الأبعاد أداء متميزاً باستمرار وهذه الخوارزميات التي أظهرت أداءً جيداً هي الخوارزميات المعتمدةعلى تحليل المكونات الأساسيةprincipal component analysis(PCA) وتحليل التمييز الخطي linear disciminent analysis(LDA)

وتحليل الملامح المحلية local feature analysis(LFA)

وتطابق الشكل والمجموعة المرنelastic punch graph matching (EBGM).

كل هذه الخوارزميات كان أداؤها جيداً من أجل صور الوجوه ثنائية الأبعاد الأمامية والقريبة من الأمامية والملتقطة داخل غرف لكن أداءها ينخفض مع تغير ظروف الإضاءة والدوران في الرأس .

هذه الخوارزميات الجيدة للتعرف على الوجوه ثنائية الأبعاد شجعت على تطوير خوارزميات مماثلة للتعرف على الوجوه ثلاثية الأبعاد. أكثر من ذلك، كثير من هذه التقنيات ثنائية الأبعاد تم جمعها مع تقنيات ثلاثية الأبعاد لتشكيل خوارزميات ثنائية+ثلاثية الأبعاد (٢D+٣D) .

طالما أن التركيز في هذا البحث لن يكون على التعرف ثنائي الأبعاد على الوجه فإن الخوارزميات الرئيسية فقط ستتم مناقشتها بما أنها تشكل القاعدة لكثير من استرتيجيات ٣D و٢D+٣D .

١-٣-١ التقنيات المترابطة المعتمدة على المظهر appearance based techniques :

من بين التقنيات المترابطة للتعرف على الوجوه ثنائية الأبعاد المعتمدة على مظهر صور الوجه الرمادية أو الملونة، هناك عدد من طرق الإسقاط في فضاء جزئي. وتضم خوارزميات تستخدم تحليل المكونات الأساسية(PCA) والتحليل التمييزي الخطي(LDA). في هذه الطرق يتم التعبير عن صورة وجه معين كنقطة في فضاء خواص ذو N بعد حيث N عدد البكسلات في الصورة. كل صور الوجه البشرية يتم نمذجتها لتتوضع في فضاء جزئي خطيأغير خطي من فضاء الخواص ذو الN بعد. هناك طرق احصائية تستخدم لتعليم الفضاء الجزئي باستخدام مجموعة من صور الوجوه. يتم اسقاط كل صور الوجوه في الفضاء الجزئي قبل القيام بعملية التصنيف.

الطرق المعتمدة على المظهر المترابط تستخدم عادة معلومات من كامل الصورة ولا تحتاج إلى معرفة عالية المستوى عن وجوه الناس للحصول على أجزاء من الوجه وبالتالي فهي أقل قوة تجاه تغير خلفيات الصور واقتطاع مناطق من الوجه. أكثر من ذلك بهدف تعلم الفضاء الجزئي للوجوه بشكل موثوق أكثر فإنها تحتاج عدداً كبيراً من صور التدريب لعدة أشخاص وملتقطة تحت ظروف متنوعة. سوف تتم مناقشة تقنيتي ترابط مهمتين للتعرف على الوجوه ثنائية الأبعاد هما PCA و LDA .

١-٣-١-١ تحليل المكونات الأساسية principal component analysis:

PCA أو الوجوه الخاصة كانت واحدة من أول التقنيات الناجحة من أجل التعرف ثنائي الأبعاد على الوجه. في هذه التقنية يتم الحصول على مجموعة من الأشعة المتعامدة، والتي يكون التغير في بيانات صور الوجوه أعظمياً وفقها، بواسطة التحليل الخاص لمصفوفة التغاير لصور الوجوه وتسمى الأشعة الخاصة بالوجوه الخاصة.

ينتج عن PCA تمثيلات مخفضة لبيانات ذات أبعاد كبيرة وهذا مثالي من أجل متوسط المربع أي أن PCA تصل إلى أقل قيمة لمتوسط مربع الخطأ LMS least mean squared بين الصورة الأصلية والصورة المقابلة المعاد تشكيلها من المحاور المحددة بالأشعة الخاصة. إذا تم إعادة تشكيل صورة معينة من كل المحاور الخاصة فإن متوسط مربع الخطأ بين الصورة الأصلية ونسختها المعاد تشكيلها يساوي الصفر .

من أجل التعرف على الوجوه فإنه يتم الاحتفاظ بالأشعة الخاصة M الأكبر من بين الأشعة الخاصة الكلية N (M≤N) والتي تمثل معظم التغير في البيانات. كل صور الوجوه تحت الاختبار وصور الوجوه في قاعدة البيانات يتم اسقاطها على M محور. يتم مقارنة الوجوه في هذا الفضاء الجديد بواسطة مقياس مسافة مناسب.

PCA مناسبة من حيث أنها حل مغلق ونتج عن استخدامها خوارزميات فعالة للتعرف على الوجوه ثنائية الأبعاد وتعتبر معياراً يتم مقارنة خوارزميات أخرى معها. على الرغم من نجاحها في التعرف على الوجوه ثنائية الأبعاد فإنه ينخفض أداؤها مع التغيرات في الوجه مثل تغيرات التعابير والوضعية والاضاءة.

١-٣-١-٢ التحليل التمييزي الخطي linear discriminate analyses :

هناك طريقة أخرى ناجحة للتعرف على الوجوه ثنائية الأبعاد وهي ترتكز على تحليل التمييز الخطي لفيشر وتسمى هذه الطريقة أيضاً وجوه فيشر fisherfaces. في هذه التقنية يتم اسقاط البيانات عالية الأبعاد بشكل خطي على C-١ LDA محور. حيث C عدد الأصناف حيث أن نسبة التوزع بين (خارج) الأصناف إلى التوزع ضمن الأصناف يتم رفعها بأكبر ما يمكن. يمكن لLDA أن يميز بنجاح بين أصناف قابلة للفصل خطياً. من أجل التعرف على الوجوه يتم أولاً خفض أبعاد البيانات بواسطة PCA أو تقنية أخرى قبل تطبيق LDA وذلك للتأكد من أن مصفوفة التوزع داخل الأصناف، والتي تدخل في حسابات LDA، ليست مفردة.

من أجل التعرف على الوجوه تم الحصول على نتائج أفضل باستخدام LDA منه باستخدام PCA. يمكن شرح ذلك بالحقيقة التي تقول أنه بينما يسقط LDA البيانات إلى المحاور الجديدة بحيث تكون الأصناف منفصلة لأكبر حد فإن PCA غير ملائم بالنسبة لمشاكل التصنيف .

١-٣-٢ التقنيات المعتمدة على الملامح المحلية(local features) :

خوارزميات التعرف على الوجوه ثنائية الأبعاد المعتمدة على قوالب لملامح\أبعاد الوجه المحلية تستخدم خواص مناطق محددة من الوجه كخواص من أجل التعرف وبالتالي تحتاج مثل هذه التقنيات إلى خطوة إضافية وهي التحديد الأتوماتيكي لأجزاء محددة من الوجه باستخدام قوالب أبعاد مرنة أو خواص الكثافة intensity\البنية لملامح محددة من الوجه.

وبالتالي يعتمد أداء الطرق المعتمدة على الملامح على التحديد الدقيق لنقاط العلام في الوجه فعملية التقسيم للحصول على ملامح الوجه مسألة مهمة وأساسية. يجب أن تكون تقنيات التقسيم محددة كفاية لكي تحدد فقط ملامح الوجه المطلوبة ولكنها أيضاً عامة كفاية لكي تحدد هذه الملامح من أجل صور للوجوه متنوعة ومختلفة. هذان الهدفان متناقضان وبالتالي من الصعب تحقيقهما معاً.

إذا كان بالامكان تحديد ملامح الوجه بشكل موثوق فإنه يمكن استخدام تقنيات فعالة للتعرف على الوجوه معتمدة على ملامح الوجه المحلية. وتكون هذه التقنيات قوية أكثر تجاه التغيرات في وضعية الوجه والتعابير والاضاءة والحفر وتغطية بعض مناطق الوجه ووجود الضجيج أكثر من الطرق المعتمدة على المظهر الكلي وذلك لأن بعض الملامح الوجهية التي نحصل عليها من مناطق الوجه المحددة والمقتطعة قد تبقى بدون تغيير عند تغير الظروف الأخرى.

أكثر من ذلك تكون المعلومات المرمزة في تقنيات التعرف على الوجوه المرتكزة على ملامح الوجه المحلية أسهل تفسيراً فمثلاً قدرة التمييز بين أجزاء الوجه المتنوعة يمكن أن تقيم بشكل مستقل عن الأجزاء الأخرى.

فيما يلي شرح عن تقنيتين ناجحتين للتعرف على الوجوه ثنائية الأبعاد معتمدتين على خواص الوجه المحلية وتسميان LFA local feature analysis و EBGM elastic bunch graph matching .

١-٣-٢-١ تحليل الملامح المحلية local feature analyses:

PCA لا تستغل الترابطات بين البكسلات المتجاورة في صورة. المحاور الخاصة جداً ليست محاور طبوغرافية بمعنى أنه من غير المفهوم كيف ترتبط الوجوه الخاصة مع بعضها البعض. للتغلب على بعض هذه المشاكل في PCA فإن طريقة تحليل الملامح المحلية تم تطويرها لتستخدم من أجل التعرف على الوجوه.

باعتبارأن برنامج هذه الخوارزمية يعتبر برنامج خاص بشركةIdentix R,Minnetonka, MN فإن التفاصيل الخاصة بكيفية عمل هذه التقنية غير متوفرة. لكن من المعروف أن LFA محلية بمعنى أن مكونات LFA تلتقط التغيرات في المناطق الجزئية من الوجه. لكن كما هو الحال في PCA و LDA يتم تطبيقها مباشرة على كل مناطق الوجه بدون الخطوة الأولية (تقسيم الوجه إلى أجزاء مختلفة). أجزاء LFA هي تمثيل للوجه محلي متوزع طبوغرافي تم الحصول عليه بطريقة احصائية. هذا التمثيل يمكن أن يعطي معلومات عن امكانية التمييز والتفريق لمختلف أجزاء الوجه. LFA تؤدي جيداً من أجل صور الوجوه الأمامية ثنائية الأبعاد الملتقطة تحت اضاءة ثابتة. المرجع[٢]

١-٣-٢-٢ تطابق الشكل والمجموعة المرن(elastic Bunch graph matching (EBGM:

طور العالم Wiskott تقنية أخرى ناجحة من أجل التعرف على الوجوه ثنائي الأبعاد تسمى تطابق الشكل والمجموعة المرن. EBGM مرتكزة على معاملات مرشح Gabor المستنتجة من نقاط محددة موثوقة من الوجه. في هذه التقنية يتم تحديد نقاط موثوقة يدوياً في صور قاعدة البيانات. من كل نقطة يتم الحصول على jet مؤلف من ٤٠ معامل من أجل مرشحات Gabor عند ٥ ترددات فراغية و٨ اتجاهات. يتم ربط ال jets من كل صور قاعدة البيانات لتشكيل بنية معطيات تسمى "bunch-مجموعة". أيضاً يتم تشكيل شكل وجه مرن بربط نقاط موثوقة بخطوط مستقيمة. شكل الوجه والمجموعة يشكلان معاً بنية معطيات تسمى "الشكل والمجموعة المرنين" EBG. تستخدم EBGs من أجل كل من التحديد الأتوماتيكي للنقاط الموثوقة في الوجوه تحت الاختبار ومن أجل التعرف على الوجه. تعمل EBGM بشكل جيد من أجل تعابير الوجه المتغيرة وظروف الاضاءة المتغيرة لكن تعمل بشكل سيء من أجل التغيرات الكبيرة في وضعية الرأس .

١-٤ نماذج الوجوه ثلاثية الأبعاد:

في هذا القسم سنشرح التقنيات المستخدمة للحصول على نماذج وجوه ثلاثية الأبعاد في الزمن الحقيقي وكذلك بنية البيانات المستخدمة لتمثيل هذه النماذج من أجل خوارزميات التعرف على الوجوه ثلاثية الأبعاد .

١-٤-١ التقاط نموذج الوجه ثلاثي الأبعاد :

يتم التقاط نماذج الوجه ثلاثية الأبعاد باستخدام تقنيات فعالة وغير فعالة. قدم بيسل تلخيصاً لتقنيات التقاط الصور ثلاثية الأبعاد. تقنية الالتقاط الفعالة ثلاثية الأبعاد الأكثر استخداماً مرتكزة على مواسح فراغية ليزرية. الماسح الليزري يسقط ضوءاً من المنبع الليزري على مشهد معين ويسجل انعكاسه وعمق السطح الأقرب إلى الكاميرا يتم تحديده بواسطةtriangulation. الماسحات الليزرية تعطي نماذج ثلاثية الأبعاد دقيقة وكثيفة لكنها تحتاج وقتاً أطول للالتقاط من التقنيات غير الفعالة وأكثر من ذلك تحتاج هذه الماسحات أن يكون الانسان ثابتاً تماماً خلال عملية التقاط الصورة وبالتالي فإن الماسحات الفراغية الليزرية غير مناسبة من أجل التطبيقات التي يجب أن يكون فيها المسح سريعاً وهناك مشكلة أخرى وهي الطبيعة المزعجة لضوء الليزر بالنسبة لعين الانسان .

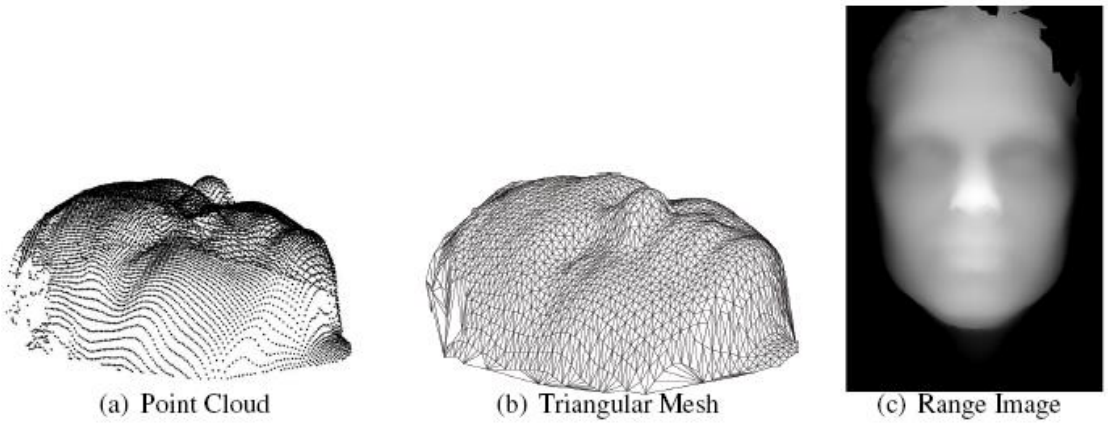
التقنيات غير الفعالة المستخدمة لالتقاط صور الوجوه ثلاثية الأبعاد تضم تصوير stereo والطرق المعتمدة على الضوء المركب. في أنظمة تصوير الstereo تقوم عدة كاميرات باستمرار بالتقاط صور الوجه من عدة زوايا. يتم الحصول على معلومات العمق باستخدام بارامترات ضبط الكاميرا ومعلومات الاختلاف التي نحصل عليها من زوايا الرؤية المختلفة.

في طريقة الضوء المركب يتم اسقاط نموذج ضوء نظامي (مثلاً نموذج خط ضوء) على المشهد. يتم تغيير نموذج الضوء و بارامترات ضبط الكاميرا لمعرفة عمق كل نقطة في المشهد بواسطة عملية triangulation . بينما تكون التقنيات غير الفعالة أسرع وأكثر أمناً وأرخص من الماسحات الليزرية فإنها أقل دقة وفيها بيانات مفقودة أكثر .

جرت أيضاً محاولات للحصول على شكل ثلاثي الأبعاد للوجه من صورة أو أكثر ثنائية الأبعاد وذلك بتعديل نماذج عامة لوجوه ثلاثية الأبعاد. العديد من هذه التقنيات تتضمن تعديلاً على خوارزميات "شكل من X " لكن تشكيل شكل ثلاثي الأبعاد من صورة ملامح واحدة عملية صعبة. مثل هذه التقنيات يمكن أن تكون مفيدة لتوليد نماذج وجوه ثلاثية الأبعاد للأشخاص الذين تكون صورهم ثنائية الأبعاد متوفرة لكن لا يمكن التقاط صورهم ثلاثية الأبعاد في الزمن الحقيقي.

١-٤-٢ تمثيل سطح الوجه :

غيوم النقاط أو شبكات السطح المثلثية أو الصور الفراغية جميعها تستخدم لتمثيل سطوح الوجوه. تمثيل غيمة النقاط تحوي الإحداثيات (x,y,z) لمجموعة من النقاط على سطح الوجه (الشكل a ١-٤). يمكن وصل هذه النقاط إلى جيرانها الأقرب بواسطة خطوط مستقيمة مما ينتج عنه تمثيل شبكي مثلثي (الشكل b ٤). يمكن ضغط تمثيل غيمة النقاط أو التمثيل الشبكي المثلثي بأخذ عينات بكثافة من السطوح التي تحوي معلومات مفصلة وعينات متفرقة من المناطق الملساء نسبياً. النقاط ثلاثية الأبعاد في طرق التمثيل هذه عادة غير مرتبة وبالتالي تحتاج خوارزميات معقدة نوعاً ما من أجل معالجتها.



الشكل(١-٤)الصورة الفراغية الشبكة المثلثية غيمة النقاط

الصورة الفراغية والتي تسمى أيضاً سطح ٢,٥ بعد (٢,٥D) أو خريطة العمق تتألف من نقاط (x,y) في شبكة مستطيلة نظامية. كل نقطة (x,y) ترتبط بقيمة z للنقطة على السطح الأقرب إلى جهاز الالتقاط (الشكل c ١-٤). يمكن الحصول على الصور الفراغية بواسطة اسقاط متعامد لشبكات السطوح أو غيوم النقاط ثلاثية الأبعاد. أيضاً تتوفر أجهزة الالتقاط ثلاثية الأبعاد والتي تلتقط صور فراغية مباشرة.

عند وضع النقاط في صورة فراغية في شبكة نظامية وفق مستطيلات فإنه يمكن معالجتها بواسطة خوارزميات معالجة صور بسيطة نوعاً ما.

بالمقارنة تكون تقنيات التعرف على الوجوه ثلاثية الأبعاد أفضل من التقنيات ثنائية الأبعاد حيث يمكن وبسهولة نسبياً تصحيح وضعية النماذج ثلاثية الأبعاد بواسطة تدوير rotation و ازاحة translation وهذا يحل كثيراً مشكلة الوضعية.

عادة تكون النماذج ثلاثية الأبعاد في محاور العالم الحقيقي وبالتالي فهي لا تحتاج إلى إعادة تقييس قبل المعالجة. يمكن القول نظرياً أن شكل الوجه مستقل عن العوامل الخارجية مثل ضوء البيئة المحيطة أثناء التقاط الصورة. إن معرفة شكل الوجه واتجاه وكثافة الضوء المحيط أثناء الالتقاط يمكننا من حساب خواص الانعكاس عن سطح الوجه. يمكن تطبيق ذلك لتوليد صور للوجه بشكل مركب تحت ظروف اضاءة مختلفة والتي تساعد أيضاً في تخفيف مشكلة اضاءة الوجه.

هناك مشاكل تتعلق بالطبيعة المتغيرة للإضاءة في خوارزميات التعرف على الوجوه ثلاثية الأبعاد وبما أن صور الكثافة تستخدم لتشكيل نماذج ثلاثية الأبعاد في التقنيات غير الفعالة للحصول على الصور ثلاثية الأبعاد فإن التغيرات في الإضاءة يمكن أن تغير شكل النموذج المركب. لكن وفي دراسة مؤخراً ظهر أن أداء خوارزميات التعرف على الوجوه ثلاثية الأبعاد لا تتغير كثيراً عندما يتم التقاط النماذج ثلاثية الأبعاد تحت ظروف اضاءة متغيرة باستخدام جهاز تصوير فراغي. ولوحظ أيضاً أنه بينما يتحسن أداء خوارزميات التعرف على الوجوه ثنائية الأبعاد بشكل كبير عندما يتم تعويض تغيرات الاضاءة فإنه لا يوجد تحسن ملحوظ في أداء الخوارزميات ٢D+٣D. وهذا الدليل يثبت أن خوارزميات التعرف على الوجوه ثلاثية الأبعاد يمكن أن تتأثر أقل بتغيرات ظروف الإضاءة من خوارزميات ٢D .

١-٥ طرق التعرف على الوجوه ثلاثية الأبعاد 3D face recognition techniques :

الصنفان الأساسيان لطرق التعرف على الوجوه ثلاثية الأبعاد هما :

أ- الطرق المعتمدة على مظهر الصور الفراغية للوجوه.

ب-الطرق المعتمدة على الخواص الهندسية لسطوح الوجوه ذات "النموذج الحر free form"

على الرغم من أنه تمت دراسة عدد من خوارزميات التعرف على الوجوه ثلاثية الأبعاد و٢D+٣D فإنه تم تقييم معظمها من أجل قواعد بيانات صغيرة وخاصة غير متاحة لبقية الباحثين. هناك بعض الدراسات التي وضحت أداء عدة خوارزميات تعرف على الوجوه ثلاثية الأبعاد على قاعدة بيانات مشتركة.

تختلف دراسات التعرف على الوجوه ٣D أو ٢D+٣D عن بعضها من حيث التصميم والبوتوكول المستخدم لتقييم الدراسة. فمثلاُ عملت مجموعات بحث مختلفة على قواعد بيانات مختلفة الحجم وبالتالي لتفسير نتائجها يجب الأخذ بعين الاعتبار أن أداء خوارزميات التعرف عل الوجوه ينخفض مع زيادة عدد الأشخاص في قاعدة البيانات وكذلك تم استخدام أعداد مختلفة من الصور لكل شخص في بناء قاعدة البيانات في دراسات مختلفة ويجب الاشارة إلى أن زيادة عدد الصور لكل شخص في قاعدة البيانات يحسن عادة من أداء خوارزميات التعرف.

هناك أيضاً عدم انسجام في طريقة عرض النتائج في الأبحاث المختلفة فمثلاً من أجل تقييم أداء عملية التعرف تم استخدام قيمة معدل التعرف RR من المرتبة n بدلاً من قيمة معدل التعرف من المرتبة ١. وكذلك من أجل تقييم أداء معدل التحقق تم أحياناً استخدام أحد القيمتين FAR أو FRR من أجل نقطة عمل معينة بدلاً من استخدام كلاً من القيمتين. وبالتالي من الضروري اختبار خوارزميات التعرف على الوجوه ٢D و ٣D من أجل قاعدة بيانات كبيرة مشتركة باستخدام بروتوكول تقييم موحد ثم يتم مقارنة الخوارزميات.

وفي هذا الاتجاه تم بناء قاعدتي بيانات كبيرتين هما FRGC v٠,١ و FRGC v٠,٢ وتم تقييم أداء بعض الخوارزميات من أجلهما بناء على بروتوكول موحد. وتم اعتبار أداء خوارزميات ٣D PCA كمرجع للمقارنة وتم مقارنة كل الخوارزميات الأخرى بالنسبة لها.

١-٥-١ طرق التعرف على الوجوه ثلاثية الأبعاد المعتمدة على مظهر الصور الفراغية للوجوه "appearance based":

تشبه طرق التعرف المعتمدة على المظهر المترابط ثنائي الأبعاد والاختلاف الوحيد عنها أنها تستخدم صور فراغية range images عوضاً عن صور الكثافة intinisity images وفي معظمها تشكل امتداداً عن التقنيات التي أثبتت نجاحها مع صور الوجوه ثنائية الأبعاد.

هناك حاجة لعدد من خطوات المعالجة الأولية والتقييس normalization في هذه الخوارزميات والهدف من هذه الخطوات هو تحديد واقتطاع رأس الانسان وازالة النتوءات الناتجة عن الضجيج وملئ الفراغات (المناطق الناتجة عن بيانات ناقصة مفقودة) وكذلك وضع الوجوه في موضع نظامي موحد وتوليد صور فراغية في ذلك الموضع. يتم تحديد ثلاث نقاط أو أكثر في الوجه أتوماتيكياً أو يدوياً لتحديد وضعية الرأس ومن أجل معظم الخوارزميات يكون الوضع النظامي هو الوضعية الأمامية للوجه مع وجود قمة الأنف في مركز الصورة.

الطرق المعتمدة على المظهر التي تم البحث فيها من أجل التعرف على الوجوه ثلاثية الأبعاد تضم PCA,LDA وكذلك تحليل المكونات المستقلة(ICA) ونماذج ماركوف المخفية(HMM) وتحليل المكونات المثالية(OCA).

١-٥-١-١ التحليل إلى المكونات الأساسية PCA:

تعتبر PCA الطريقة المعتمدة على المظهر الأكثر بحثاً وتعتبر مرجعاً لتقييم أداء خوارزميات التعرف على الوجوه ثلاثية الأبعاد الأخرى[الجداول(١-١)،(١-٢)]. وتم استخدام هذه الطريقة أيضاً في مجال الكشف ثلاثي الأبعاد عن الوجه لفصل مناطق الوجه عن بقية أجزاء الصورة الفراغية[١١]. يتم في هذه الطريقة حساب السطوح الخاصة بشكل مشابه لحساب الوجوه الخاصة ثنائية الأبعاد. وجد أن resolution الفراغية المناسبة من أجل ٣D PCA تساوي ٠,٩٨ mm لكل بكسل على طول المحورين x,y وتتراوح بين ٠,٥ to ٣ mm لكل بكسل على طول المحور z [١٠]. وهذا يعني أن كل resolution أكبر من ذلك غير لازم.

واجمالاً تتراوح معدلات التعرف من المرتبة الأولى بين ١٠٠% و ٦٨% بواسطة 3D PCA وذلك من أجل قواعد بيانات مختلفة الحجم والتعقيد [الجداول (١-١)-(١-٢)]. على الرغم من أن التجارب الأولية على قواعد بيانات أصغر أعطت أداء جيداً من أجل 3D PCA فإن الأداء الأكثر واقعية يمكن الحصول عليه من الدراسات الأخيرة التي أجريت على قواعد بيانات أكبر وأكثر تعقيداً.

تم تطبيق PCA على قيم العمق للصور الفراغية[١٠-١٢] أو على القيم (x,y,z) لغيمة نقاط الوجه ثلاثية الأبعاد[١٤] أو على تمثيل كروي نظامي لسطوح الوجوه[١٣] وبطرق أخرى لكن من غير الواضح أيها أكثر كفاءة للاستخدام من أجل PCA. أظهر أحد الباحثين Pan أن تطبيق PCA على تمثيل كروي نظامي لسطوح الوجوه بدلاً من تطبيقها مباشرة على الصور الفراغية يعطي نتائج أفضل [١٣] حيث تم مقارنة الأسلوبين من أجل قاعدة البيانات FRGC v٠,١.

وكذلك تم تطبيق PCA على الصور الفراغية كاملة[١٢] أو فقط على المناطق من الصور الفراغية التي تحتوي على ملامح الوجه الرئيسية [١٥-١٠-١٤]. من المحتمل أن استبعاد المناطق من الصورة الفراغية التي لا تعتبر من ملامح الوجه الرئيسية قد يكون مفيداً حيث أنه قد يزيل الضجيج غير المرغوب بسبب موضات الشعر أو اللباس.

تم مقارنة أداء PCA مع الصور ذات التدرج الرمادي ثنائية الأبعاد مع 3D PCA في عدة دراسات[١٦-١٠]. لكن لم يتم الفصل في أي منهما أفضل من الآخر حيث ظهر في دراسات أنهما متقاربان في الأداء [١٧] وفي دراسات أخرى ظهر أن 3D PCA أفضل من 2D PCA [١٠] ولكن في دراسات أخرى ظهر أن استخدام الصور الملونة بدل الصور ذات التدرج الرمادي في 2D PCA يعطي نتائج أفضل من 3D PCA [١٨].

لوحظ في كل هذه الدراسات أن الجمع بين أسلوبي التعرف 2D مع 3D ينتج عنه أداء أفضل مقارنة مع كل أسلوب على حدى.

١-٥-١-٢ تحليل المكونات المستقلة ICA:

لا تأخذ ICA بعين الاعتبار فقط العلاقات الخطية بين البكسلات في صورة فراغية ولكن أيضاً العلاقات من مرتبة أعلى فهي تسقط البيانات بشكل خطي على محاور جديدة مستقة احصائياً قدر ما يمكن[١٩]. وبشكل مماثل للنتائج التي تم الحصول عليها من أجل التعرف على الوجوه ثنائية الأبعاد فإن أداء ٣D ICA كان أفضل من ٣D PCA [٢٠].

١-٥-١-٣ التحليل التمييزي الخطي LDA:

تم البحث في تقنيات التحليل التمييزي الخطي من أجل التعرف على الوجوه ثلاثية الأبعاد[الجداول (١-١)،(١-٣)]. تم تطبيق LDA على أنواع تمثيل ثلاثية الأبعاد متنوعة قد تكون الصور الفراغية أو ميول أفقية وشاقولية للصور الفراغية. أعطت طرق التمثيل بواسطة الميول أفضل أداء للتعرف [٢١] لكنه ليس واضحاً أيهما أفضل الميول الأفقية أو الشاقولية.

وأظهرت النتائج أن أداء ٣D LDA أفضل من ٣D PCA [٢٢] وكذلك أفضل من ٢D LDA وأن ٢D+٣D LDA أفضل من الأسلوبين(٢D أو ٣D) على حدى [٢١].

١-٥-١-٤ تحليل المكونات المثالية OCA:

بحث Srivastava في تقنية تسمى تحليل المكونات المثالية من أجل التعرف على الوجوه ثلاثية الأبعاد [٢٣] حيث تم الحصول على المساقط الخطية المثالية للبيانات في فضاء ذو أبعاد أقل بحيث يتم الحصول على أكبر أداء لمصنف النقطة الأقرب. تتضمن الخوارزمية عملية تكرارية مكلفة حسابياً لتعلم الأشعة الأساسية. ويقول الباحثون أن هذه الطريقة أفضل من PCA,LDA,ICA من أجل التعرف على الوجوه ثلاثية الأبعاد وكانت قادرة على التعامل مع تغيرات في التعابير.

١-٥-١-٦ نماذج ماركوف المخفية HMM:

تم تطبيق تقنيات التعرف على الوجوه ثنائية الأبعاد المعتمدة على نماذج ماركوف المخفية(HMM) من أجل الصور الفراغية للتعرف على الوجوه ثلاثية الأبعاد [١٢]. تستفيد هذه التقنيات من حقيقة أن ملامح الوجه تأتي بشكل طبيعي وفق ترتيب ثابت من الأعلى إلى الأسفل ومن اليسار إلى اليمين بغض النظر عن التغيرات في الاضاءة أو الوضعية أو تعابير الوجه. تتم نمذجة مكونات الوجه كحالات في نماذج ماركوف التي يتم تعلمها من مجموعة من صور الوجوه. تم دراسة الطرق المعتمدة على نماذج ماركوف المخفية المدمجة embedded hidden marcov models من أجل التعرف على الوجوه باستخدام الصور الملونة أو صور التدرج الرمادي أو الصور الفراغية أو من أجل جمع بين هذه الأساليب -٢٤] [٢٧-٢٦-٢٥. تقنية ٣D EHMM كان أداؤها أقل من EHMM مع الصور ذات التدرج الرمادي وكذلك أقل من أداء تقنية EHMM مع الصور الملونة [٢٧-٢٦]. ومثل باقي الطرق المعتمدة على المظهر فإن جمع الأساليب الثلاثة الأخيرة يعطي أفضل من كل أسلوب على حدى.

١-٥-٢ تقنيات التعرف على الوجوه المعتمدة على "النموذج الحر free form":

يشار إلى العناصر ثلاثية الأبعاد التي لا يمكن اعتبارها مسطحة أو سطح مربع بطبيعته ب"النموذج الحر" ويعتبر وجه الانسان مثالاً عن "النموذج الحر" [٢٨].

١-٥-٢-١ مقارنة سطح الوجه surface matching:

تم تطوير عدة تقنيات للتعرف على الوجوه ثلاثية الأبعاد تم فيها مقارنة سطحي وجهين مباشرة [الجداول(١-٣)،(١-٤)،(١-٦)]. الفكرة الرئيسية لهذه الطرق هي مطابقة وضعيتي pose سطحي الوجهين وموضعهما position بأقرب ما يمكن في الفضاء ثلاثي الأبعاد ثم مقارنتهما بواسطة مقياس مناسب. تم مناقشة عدة أساليب لهذه التقنية الأساسية بتغيير اجراء مطابقة الوضعية والموضع أو تغيير المقياس المستخدم لحساب المسافة بين السطحين.

هناك أسلوبان أساسيان لمطابقة وضعيات وموضع سطوح الوجوه هما التقييس غير الدقيق والتقييس غير الدقيق المتبوع بالتقييس الدقيق. في التقييس غير الدقيق [٣٠-٢٩] يتم حساب المكان والوضعية لكل سطح وجه في الفضاء ثلاثي الأبعاد ويتم تحويلها إلى وضعيات أمامية نظامية ثم يتم مقارنة السطحين وهما في هذا الوضع بواسطة مقياس متوسط مربع الخطأ (MSE) [٣٠] أو مسافة (HD)Hausdroff [٢٩] التي قدمها Huttenlocher. وهذا الأسلوب في التقييس غير كافي لأن مقاييس المسافة هذه تعتمد على تطابق الموضع والوضعية الدقيق للسطحين.

لذلك في معظم طرق التعرف على الوجوه ثلاثية الأبعاد التي تستخدم مقارنة سطوح الوجوه فإنه يتم اتباع مرحلة التقييس غير الدقيق بمرحلة التقييس الدقيق [٣١-٣٢] والأسلوب الأكثر استخداماً من أجل التقييس الدقيق هو خوارزمية النقطة الأقرب التكرارية(ICP). في ICP يتم تحويل أحد السطحين ثلاثيي الأبعاد بخطوات متكررة في الفضاء ثلاثي الأبعاد وذلك لوضعه أقرب ما يمكن إلى السطح الآخر من حيث المكان والوضعية. والتقييس غير الدقيق يساعد اجرائية التقييس الدقيق التكرارية لتصل إلى مطابقة موضع السطحين ووضعيتهما بشكل صحيح وتمنعها من الوقوع في مشكلة النهاية المحلية.

استخدمت عدة أبحاث متوسط مربع الخطأ بين أزواج النقاط الأقرب في السطحين على أنه المعيار الذي يجب تصغيره في الاجرائية المتكررة [٣٢] وفي دراسات أخرى تم استخدام مقياس HD أو HD الجزئية [٣٣] أو جمع بين المقياسين MSE و HD [٣٤] أو جمع بين MSE بين النقاط الأقرب مع MSE بين السطوح والنقاط الأقرب [٣٣] كما قدم ذلك Chen و Medioni [٣٥]. وثبت أيضاً أن استخدام MSE المعتمدة على المسافات بين النقاط والسطوح [٣٥] في تقنية ICP بدلاً من MSE المعتمدة على المسافات بين نقطة ونقطة [٣٦] يقلل من احتمال وقوع هذه التقنية في النهاية المحلية. انخفض الأداء قليلاً من أجل نفس قاعدة البيانات عند استخدام نقاط عشوائية من منطقة الأنف فقط من أجل عملية مقارنة سطوح الوجوه[٣٤].

أعطت ICP وتعديلاتها تقنيات تعرف ناجحة من أجل التعرفD٣ و ٢D+٣D [الجداول(١-٣)، (١-٤)،(١-٦)]. أظهرت تقنيات التعرف المعتمدة على ICP قوة تجاه تغيرات وضعية الوجه[٣٨] وتغير ظروف الإضاءة خلال الحصول على الصورة ثلاثية الأبعاد[٣٩] وأداؤها أفضل من طرقPCA ٣D [٣١]وLFAمع الصورالملونة[٤١] و LDA ٢D [٣٨].

من غير الواضح فيما إذا كان اضافة معلومات الصورة الملونة أو الرمادية ثنائية الأبعاد إلى تقنية ٣D ICP قد يؤدي زيادة في أدائها. حصل Lu على أداء أفضل عند جمع ٢D LDA+٣D ICP مقارنة مع أداء كل تقنية على حدى[٣٨] وكذلك لاحظ Maurer تحسن أداء ٣D ICP عند جمعها مع ٢D EBGM [٤٢]. في حين لم يلاحظ تحسن كبير في الأداء عند استخدام معلومات الصورة ذات التدرج الرمادي مع قيم الاحداثيات (x,y,z) للحصول على ٤D ICP مقارنة مع حالة استخدام الاحداثيات الفراغية فقط[٤٣].

١-٥-٢-٢ مطابقة المنحنيات الجانبية profile matching:

يمكن الحصول بسهولة على منحنيات جانبية وحيدة البعد من نماذج الوجوه ثلاثية الأبعاد بقطع هذه الوجوه بمستويات بتوجهات مختلفة. هناك عدد من الدراسات التي تم فيها مقارنة أشكال منحنيات جانبية مختلفة للوجه بهدف التعرف على الوجوه[الجداول(١-٣)،(١-٥)،(١-٦)] أظهرت عدة دراسات أن المنحني الجانبي في المستوي الشاقولي المركزي للوجه هو الذي يحوي أكبر قدر من المعلومات المفيدة للتمييز بين الوجوه. وجد Nagamine [٤٤] أن المنحنيات الجانبية الشاقولية التي تتوضع في المنطقة المركزية للوجه هي التي تحوي أكبر قدر من معلومات التمييز ويليها المنحنيات المتمركزة ٤٠mm فوق قمة الأنف. ولاحظ أن المقاطع الأفقية تحتوي أقل معلومات تمييز وحصل Beumier على نتائج مماثلة. ولوحظ انخفاض ملحوظ في أداء التعرف على الوجوه عند استخدام ثلاث مقاطع مركزية بدل ١٥ مقطع شاقولي يشمل كل سطح الوجه[٤٥].

وجد Beumier –وبشكل منسجم مع باقي تقنيات التعرف على الوجوه– أن دمج معلومات الصورة ذات التدرج الرمادي ومعلومات الشكل للمقاطع المركزية أو الطرفية الشاقولية يعطي أداء أفضل مقارنة مع كل أسلوب على حدى[٤٦].

تتميز طرق التعرف على الوجوه ثلاثية الأبعاد التي تعتمد على مقاطع الوجه بأنها تعطي معلومات التمييز المحتواة في مناطق الوجه المختلفة ولكنها تحتاج خطوة اضافية هي التحديد الدقيق للمقاطع حيث أن دقتها تعتمد على ذلك وكذلك ينخفض أداؤها بوجود التعابير في الوجوه[٤٧].

١-٥-٢-٣ الملامح المحلية للوجوه ثلاثية الأبعاد:

عدد الدراسات التي استخدمت الملامح المحلية لسطوح الوجوه ثلاثية الأبعاد اقل عدداً مقارنة مع الأبحاث التي تعتمد على المظهر التي لا تحتاج إلى تجزيء آلي لملامح الوجه(الجدول ٦)

بعض الملامح المحلية لسطوح الوجوه التي تم استخدامها للتعرف: الوضعية-خواص الانحناء لنقاط العلام في الوجه-المسافات الاقليدية ثلاثية الأبعاد-النسب بين المسافات.

طبقت تقنية تم فيها مقارنة المناطق الجزئية من الوجه (العينين-الأنف-الفم) بشكل منفصل على قاعدتي البيانات FRGC v٠,١ و FRGC v٠,٢ [٤٨] وكان أداؤها أفضل من 3D PCA. وجمعت تقنية تعرف 2D+3D تقنية EBGM المطبقة على الصور الفراغية والصور ذات التدرج الرمادي وأدت أيضاً بشكل جيد مع قاعدة البيانات FRGC v٠,٢ [٣٤].

هناك عدة محاسن تتميز بها تقنيات التعرف على الوجوه المعتمدة على ملامح الوجه المحلية على التقنيات المعتمدة على المظهر. حيث أولاً يمكن الحصول على الملامح المحلية بالاعتماد على معرفة شاملة عن التنوع البنيوي للوجوه البشرية حيث استخدمت هذه الطريقة للحصول على تقنيات فعالة للتعرف على الوجوه. وثانياً تتأثر التقنيات المعتمدة على الملامح المحلية بشكل أقل بالتغيرات الكلية في مظهر صور الوجوه بما فيها التغيرات الناتجة عن التعابير[٥٠].

لكن هذه التقنيات لم تلق اهتماماً كبيراً بالمقارنة مع التقنيات المعتمدة على المظهر الكلي وقد يكون السبب هو أن هذه التقنيات تحتاج خطوة اضافية هي التحديد الدقيق لنقاط العلام في الوجه ودقة هذا التحديد قد تؤثر بشكل كبير على مجمل الأداء وبالتالي يعطي التحديد الدقيق خوارزميات جيدة للتعرف.

١-٥-٣ الطرق المدمجة:

تم البحث في عدة تقنيات للتعرف تجمع عدة تقنيات "نموذج حر"(الجدول ٦). وفي معظمها أعطت نتائج أفضل من التقنيات المنفردة. حيث لوحظ تحسن كبير في الأداء عند الجمع بشكل هرمي بين تقنية مقارنة السطوح مع ٣D LDA [٢٢]أو ٢D LDA [٣٨]. ولاحظ Pan تحسن كبير في الأداء عند جمع تقنية مقارنة السطوح مع تقنية مقارنة المقاطع المركزية[٥٢].

Author	Method	Data Set			Performance
		# Subjects	# Gallery Images	# Probe Images	
Achermann [2]	PCA, HMM	24	5/subject	5/subject	RR = 100%
Heseltine [42]	PCA	100	40	290	EER = 12.7%
Hesher [44]	PCA	37	5/subject	1/subject	RR = 94%
Hesher [45]	ICA	37	5/subject	1/subject	RR = 97%
Heseltine [41, 43]	LDA	230	-	-	EER = 9.3%
Srivastava [102]	OCA	67	3/subject	3/subject	RR = 99%

الجدول (١-١) يلخص تقنيات التعرف على الوجوه ثلاثية الأبعاد المعتمدة على المظهر

نلاحظ من الجدول انخفاض أداء PCA مع زيادة حجم قاعدة البيانات

Author	Method		Data Set			Performance
	2D	3D	# Subjects	# Gallery Images	# Probe Images	
Tsutsumi [109]	PCA	PCA	24	9/subject	35/subject	$FAR_{2D+3D} = 4.5\%$ $FRR_{2D+3D} = 3.7\%$
Chang [21]	PCA	PCA	166	1/subject	1/subject	$RR_{2D+3D} = 92.8\%$ $RR_{2D} = 83.1\%$ $RR_{3D} = 83.7\%$
Chang [20]	PCA	PCA	200	1/subject	1/subject	$RR_{2D+3D} = 98.5\%$ $RR_{2D} = 89\%$ $RR_{3D} = 94.5\%$
Chang [24]	PCA	PCA	127	1/subject	297	$RR_{2D+3D+IR} = 100\%$ $RR_{2D} = 90.6\%$ $RR_{3D} = 91.9\%$ $RR_{IR} = 71.0\%$
Chang [23]	PCA	PCA	198	1/subject	670	$RR_{2D+3D} = 97.5\%$
Tsalakanidou [108]	color PCA	PCA	40	2/subject 3 poses	2/subject	$RR_{color+3D} = 98.8\%$ $RR_{color} = 93.8$ $RR_{3D} = 85$
Godil [35]	color PCA	PCA	200	1/subject	1/subject	$FAR_{color+3D} = 20\%$ $FRR_{color+3D} = 1\%$ $RR_{color+3D} = 82\%$ $RR_{color} = 78\%$ $RR_{3D} = 68\%$

الجدول (١-٢) يوضح دراسات استخدمت PCA للصور الفراغية مع PCA للصور الملونة أو ذات التدرج الرمادي أو صور IFونلاحظ أن الجمع الصحيح يعطي نتائج أفضل من استخدام كل أسلوب لوحده.

Author	Method		Data Set			Performance
	2D	3D	# Subjects	# Gallery Images	# Probe Images	
BenAbdelkader [6]	LFA,	LFA,	185	2/subject	2/subject	$EER_{2D+3D} = 0\%$
Tsalakanidou [106, 107]	gray EHMM	EHMM	50	5/subject	55/subject	$EER_{2D+3D} = 5.5\%$ $EER_{2D} = 9.3\%$ $EER_{3D} = 20\%$
Tsalakanidou [105]	color/gray EHMM	EHMM	20	5/subject	1402	$EER_{2D+3D} = 7.5\%$ $EER_{color} = 15.1\%$ $EER_{2D} = 13.8\%$
						$EER_{3D} = 16.7\%$
Tsalakanidou [107]	color/gray EHMM	EHMM	20	2-5/subject	2124	$EER_{2D+3D} = 7.7\%$ $EER_{2D} = 7.9\%$ $EER_{3D} = 14.8\%$ $RR_{2D+3D} = 77.3\%$
Malassiotis [66, 67]	color EHMM	EHMM	20	2-5/subject	-	$RR_{color3D} = 99.1\%$ $RR_{color} = 90.7\%$ $RR_{3D} = 96.4\%$
Lu [62]	color	ICP, shape	18	1/subject	63	$RR_{2D+3D} = 84\%$
Lu [64, 65]	LDA	ICP	200	1/subject	598	$RR_{2D+3D} = 90\%$ $RR_{2D} = 77\%$ $RR_{3D} = 86\%$
Papatheodorou [81]	4D ICP	3D ICP	62	1/subject	5/subject	$FRR_{2D+3D} = 0\%$ $FAR_{2D+3D} = 0\%$
Beumier [11]	Profiles	Profiles	100	9/subject	-	$EER_{2D+3D} = 1.2\%$

الجدول (١-٣) يوضح عدة تقنيات ٢D+٣D حيث يوضح أن جمع تقنيتين معاً أحدهما ٢D والأخرى ٣D يعطي أداء أفضل من أداء كل منهما على حدى

Author	Method	Data Set			Performance
		# Subjects	# Gallery Images	# Probe Images	
Gordon [37]	depth MSE	8	2/subject	1/subject	RR = 100%
Lao [55]	depth MSE	10	3/subject	1/subject	RR = 96%
Lee [59]	depth μ,σ MSE	35	1/subject	1/subject	Rank 10 RR = 100%
Medioni [70]	ICP	100	6/subject	1/subject	EER < 2%
Lu [61]	ICP and shape feature	18	1/subject	113	RR = 96.5%
Irfanoğlu [49]	ICP	30	2/subject	1/subject	RR = 96.7%
Achermann [1]	Partial HD	24	5/subject	5/subject	RR = 99.2%
Lee [60]	depth weighted HD	42	1/subject	1/subject	Rank 5 RR = 98%
Pan [80]	HD-ICP	30	1/subject	2/subject	EER = 3.24%
Russ [96, 97]	HD-ICP	200	1/subject	1/subject	FRR = 2% FAR = 0%
Gupta [40]	CW-SSIM	12	1/subject	29/subject	EER = 9.13% Rank 1 RR = 98.6%

الجدول (١-٤) يوضح تقنيات تعرف على الوجوه ثلاثية الأبعاد باستخدام أسلوب مقارنة السطوح وقد أعطت هذه التقنيات نتائج جيدة إلا أن مشكلتها أنها مكلفة حسابياً

Author	Method	Data Set			Perfomance
		# Subjects	# Gallery Images	# Probe Images	
Cartoux [19]	Profiles	5	1-3/subject	1/subject	RR = 100%
Pan [78]	HD Profile matching	30	1/subject	2/subject	EER = 2.22%
Nagamine [73]	Profiles	16	9/subject	1/subject	RR = 100%
Beumier [10, 12]	Profiles	120	1/subject	5/subject	EER = 4.75%
Zhang [122]	Profiles	166	166	32	RR = 96.9%

الجدول (١-٥) يوضح تقنيات التعرف المعتمدة على مطابقة مقاطع مختلفة من الوجه ونجد أن أفضلها أداء هي المعتمدة على مطابقة منحني المقطع المركزي الشاقولي

Author	Method	Data Set			Performance
		# Subjects	# Gallery Images	# Probe Images	
Gordon [38]	Curvature, Distances	8	2/subject	1/subject	RR = 100%
Moreno [72]	Curvature, Angles, Distances	60	1/subject	6/subject	RR = 78%
Lee [58]	Angles and Distances	100	–	–	RR = 96%
Gupta [39]	Euclidean and Geodesic Distances	105	1/subject	663	Rank 1 RR = 98.64% EER = 1.4%
Pan [79]	Profiles and surface matching	30	1/subject	2/subject	EER = 1.11%
Xu [119]	z values and local shape	30	5/subject	1/subject	RR = 96.1%
Gökberk [36]	LDA and Normals and Profiles	106	3/subject	2-3/subject	RR = 99.1%

الجدول (١-٦) على اليسار توضيح لخوارزميات تعرف تعتمد على ملامح الوجه المحلية

على اليمين خوارزميات تعرف تجمع بين تقنيتي تعرف ٣D والنتيجة أداء أفضل من كل تقنية على حدى

١-٦ قاعدة البيانات ثلاثية الأبعاد GavabDB المستخدمة لاختبار الخوارزمية التي نقدمها في هذا البحث للتعرف على الوجوه:

١-٦-١ مقدمة :

نظراً لصعوبة الحصول على الأجهزة اللازمة لالتقاط صور الوجوه ثلاثية الأبعاد فقد تم استخدام قاعدة البيانات GavabDB التي سنوضح فيما يلي بعض الأمور المتعلقة بها كما يبين ذلك المرجع رقم [٣].

تم بناء قاعدة بيانات صور الوجوه ثلاثية الأبعاد(GavabDB)من أجل اختبارات التعرف الأتوماتيكي على الوجوه وتطبيقات الصور الأخرى الممكنة مثل تصحيح الوضعيةpose correction. فيما يلي شرح لخواص قاعدة البيانات هذه وشرح لعملية بناء قاعدة البيانات بالتفصيل بالاضافة إلى عملية التقاط الصور وخواص الصور ومناقشة هذه الخواص بالاضافة إلى مناقشة بعض الاستنتاجات من عملية بناء قاعدة البيانات كما هو موضح في المرجع رقم [٣].

GavabDB تحتوي على ٤٢٧ صورة وجوه ثلاثية الأبعاد ل٦١ شخص (٤٥ ذكر و١٦ أنثى) وهناك ٧ صور مختلفة لكل شخص. عمر معظم الأشخاص بين ١٨ و٤٠. يتم التعبير عن كل صورة بشبكة مربعة rectangular mesh ثلاثية الأبعاد تمثل سطح الوجه. هناك لكل شخص ٢ صورة أمامية و٤ صور مدورة بدون تعابير و ٣ صور أمامية فيها تعابير.

التعرف على الوجوه له عدة تطبيقات : التعرف الشخصي – تطبيقات الأمان – تطبيقات قانونية – والطب الشرعي وغيرها. الوجه والصوت هما الخاصتان البيومتريتان الأقل تداخلاً ولزيادة قوة الأنظمة البيومترية فإن البحث في التعرف البيومتري متعدد النموذج (والذي يوظف في الوقت نفسه أكثر من خاصية بيومترية واحدة) ينمو بشكل كبير.

إن أنظمة التعرف الأتوماتيكية على الوجوه تستخدم قواعد بيانات الصور كي يتم تقييم الأنظمة بشكل فعال. وهذه الصور يجب أن يكون لها درجة تنوع معينة (تغير في الوضعيةpose أو الاضاءة أو تعابير الوجه) من أجل الصور المختلفة للشخص الواحد.

هناك بعض قواعد البيانات المنشورة والمخصصة لأغراض البحث المختلفة : التعرف الأتوماتيكي على الوجوه، تحليل تعابير الوجه، تحديد الوضعية، اقتطاع الوجه من خلفية مليئة بالعناصر ...الخ. وقاعدة بيانات صور الوجوه المستخدمة هنا تم انشاؤها لكي تستخدم في تجارب التعرف على الوجوه.

قاعدة البيانات GavabDB تحتوي صور وجوه ثلاثية الأبعاد ٣D عبارة عن سطوح للوجوه تمثل الوجوه بواسطة شبكات ثلاثية الأبعاد. كخاصية جديدة توفر قاعدة البيانات هذه ثلاثة صور لكل شخص فيها تعابير في الوجه وتحوي أيضاً على عدة تغيرات في الوضعية لكل شخص. مثل هذه التعابير في الوجه: ابتسامة – الضحكة – وتعبير عشوائي يتم اختياره من قبل الشخص وفي الحالة الأخيرة هذه(حالة التعبير العشوائي) تم السماح بتغطية الوجه باليد أو باللسان. من جهة أخرى تم حذف اللون في هذه الصور وبالتالي فإن هذه الصور عبارة تحديداً عن شبكات ثلاثية الأبعاد .

١-٦-٢ عملية الحصول على الصور ثلاثية الأبعاد :

توضح هذه الفقرة عملية الحصول على قاعدة البيانات للصور ثلاثية الأبعاد. حيث الهدف هو الحصول على قاعدة بيانات كبيرة كفاية وغنية بالتغيرات للشخص نفسه مما يسمح لنا بتقييم درجة التكيف مع هذه التغيرات بالنسبة لأنظمة التعرف على الوجوه المختبرة.

التغيرات في صور الوجوه يمكن تصنيفها إلى تغيرات داخلية وأخرى خارجية. التغيرات الداخلية تتعلق بالوجه نفسه أما التغيرات الخارجية فتتعلق بظروف التقاط صورة الوجه. وكأمثلة عن التغيرات الداخلية لدينا تعابير الوجه واللباس ووجود الشعر. وكأمثلة عن التغيرات الخارجية لدينا الخلفية، الاضاءة، اعدادت الكاميرا والمقياس في الصورة ...الخ. قاعدة البيانات التي نستخدمها فيها ثلاثة تغيرات داخلية وأربع تغيرات خارجية ترتبط بتغير الوضعية pose

١-٦-٢-١ التغيرات :

قاعدة بيانات الوجوه يجب أن تحتوي على تغيرات كافية في صور كل شخص لكي يتم اختبار عدة تقنيات مختلفة للتعرف على الوجوه وهذه التقنيات يجب أن تكون قوية من أجل هذه التغيرات في صور الوجوه. هناك ٧ صور مختلفة ملتقطة لكل شخص :

صورتان بوضعية أمامية frontal view وبدون تعابير في الوجه Neutral facial expression (اسم الصورتين i_frontal١,i_frontal٢) حيث i رقم الشخص في قاعدة البيانات.

صورة فيها دوران نحو اليمين حول المحور y وبدون تعابير (i_derecha)

صورة فيها دوران نحو اليسار حول المحور y وبدون تعابير (i_izquierda)

صورة فيها دوران نحو الأعلى حول المحور x (النظر إلى أعلى) وبدون تعابير (i_arriba)

صورة فيها دوران نحو الأسفل حول المحور x (النظر إلى أسفل) وبدون تعابير (i_abajo)

صورة أمامية بوجود ابتسامة(i_sonrisa)

صورة أمامية بوجود ضحكة(i_risa)

صورة أمامية بوجود تعبير عشوائي يتم اختياره من قبل الشخص (i_gesto)

الشكل (١-٥) يوضح الصور ثنائية الأبعاد المقابلة للصور الملتقطة للشخص في قاعدة البيانات. يتم التقاط هذه الصور بواسطة الماسح في نفس اللحظة التي يتم فيها التقاط الشبكات ثلاثية الأبعاد المقابلة. الصورتان الأولى والثانية تمثلان صور أمامية للوجه بدون تعابير في الوجه. الصورتان الثالثة والرابعة تمثلان صور جانبية يمنى ويسرى للشخص على التوالي. في هذه الصورة يكون الدوران في الوجه ٩٠ درجة تقريباً حول المحور الشاقولي في كلي الاتجاهين. الصورتان الخامسة والسادسة تمثلان دوران الوجه حول المحور الأفقي بمقدار 35^{o} تقريباً بحيث ينظر الشخص إلى الأعلى والأسفل على التوالي. الصور الملتقطة الثلاثة الأخيرة تمثل صور أمامية بوجود تعابير في الوجه مثل الابتسامة، الضحكة وتعبير عشوائي يتم اختياره من قبل الشخص. لم يتم الأخذ بعين الاعتبار التغيرات الناتجة عن وجود النظارات أو الاضاءة. المرجع [٣]

الشكل (١-٥) صور ثنائية الأبعاد تمثل التغيرات المختلفة لشخص معين في قاعدة البيانات .

الشكل (١- ٦) يوضح الصور ثلاثية الأبعاد الملتقطة للشخص نفسه في الشكل ١-٥. يوضح هذا الشكل(١-٦) الشبكة ثلاثية الأبعاد. يتم تشكيل الشبكات meshes بواسطة احداثيات vertices النقاط المتصلة مشكلة خلايا من أربع نقاط. الشكل (١-٦) يبين نفس الشبكة المدورة لتوضيح كيف تظهر تغيرات العمق في المحور z

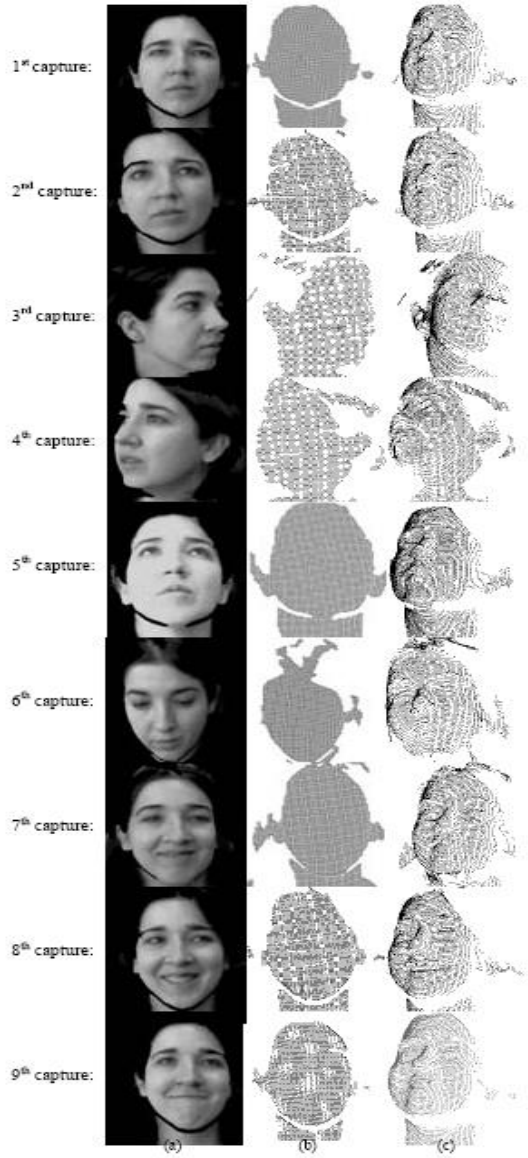
١-٦-٢-٢ بيئة الالتقاط :

فيما يلي شرح digitizer ثلاثي الأبعاد المستخدم لتشكيل قاعدة البيانات:

Minolta VI-٧٠٠ digitizer، حساس ليزري يلتقط في أقل من ثانية صورة فراغية بالاضافة إلى معلومات عن الألوان. ويولد شبكة سطح ثلاثية الأبعاد للسطح المرئي من الوجه (السطح المرئي بالنسبة للماسح هو السطح الذي يعكس الضوء الصادر عن الماسح) .

لم يتم التحكم بشروط الاضاءة خلال عملية الالتقاط. كان ضوء الشمس هو المسيطر على الرغم من أن الغرفة كانت مزودة بضوء المكاتب العادي. تم وضع الأشخاص قرب النوافذ الكبيرة .

تم وضع الوجوه على مسافة ٠,٥ m - ١,٥ من الماسح. على الرغم من أن الكرسي (مع عجلات) كان على بعد ١,٥ m فإنه كان بوسع الأشخاص تحريك رأسهم بشكل طبيعي. وهذا أدى إلى تغير في الدقة resolution من صورة لأخرى. يمكن تغيير ارتفاع الكرسي لكي يبقى الرأس مرئياً دوماً بالنسبة للماسح scanner. بهدف الحصول على توجهات الرأس المختلفة different poses فإن الأشخاص نظروا إلى علامات محددة (بدون إجبار الشخص على تثبيت وضع رأسه محافظاً على زاوية محددة بدقة)



الشكل (١-٦) :الصور ثلاثية الأبعاد للشخص الواحد a: الصور ثلاثية الأبعاد مع الألوان b :نفس الصور بدون ألوان عند دقة resolution ١/٤، c :نفس صور الوجه مدورة

كل صورة تم التقاطها في أقل من ثانية. واستغرقت عملية التقاط الصور التسعة لكل شخص حوالي عشر دقائق (بما في ذلك الوقت الأولي لشرح وضعية الشخص وتغيرات الملامح، والتحضير لكل صورة). الشكل ١-٧ يبين نظام الالتقاط المستخدم وأحد الأشخاص خلال عملية الالتقاط.المرجع [٣]

الشكل (١-٧) نظام الالتقاط للصور ثلاثية الأبعاد ووضعية الشخص بالنسبة لهذا النظام

١-٦-٢-٣ الصور ثلاثية الأبعاد الملتقطة :

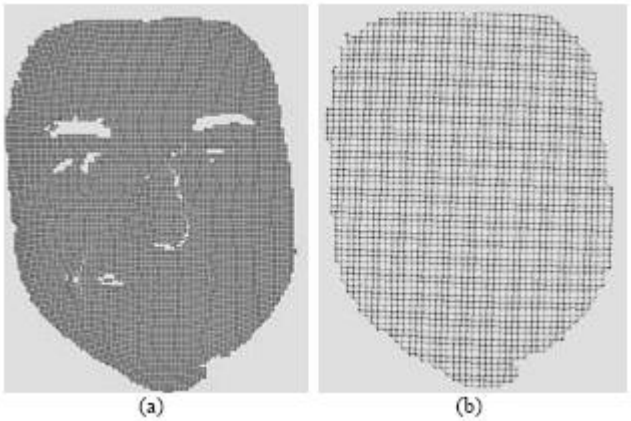
الماسح معه أداة software تسمى VIVID والتي تحوي وظائف لتخزين الصور الممسوحة إلى ملفات بأنواع مختلفة (وهذه الأنواع المستخدمة في نمذجة العناصر ثلاثية الأبعاد هي VRL,DXF,…etc. يمكن اختيار معلومات مختلفة من الصور الملتقطة لكي يتم مناقشتها :

. المجموعة الكاملة للنقاط الملتقطة (دقة resolution ١/١)أو النقاط الباقية بعد تخفيض الدقة إلى ١/١٦،١/٩،١/٤ (نقطة واحدة من أجل ٤ نقاط، ١ نقطة من أجل ٩ نقاط وهكذا). مثل هذه التخفيضات تخفض حجم الملفات بشكل كبير .

. معلومات الألوان من الشبكة mesh :يتم التخلص من الألوان(عندما لن يتم استخدامها) وذلك لتخفيض حجم تخزين صور الوجوه بشكل كبير.

النقاط غير المملوءة في الشبكة والتي قد تتواجد في مناطق لم يتم الاعتيان منها. software الماسح يوفر امكانية ملئ الفراغات الصغيرة بواسطة الخيار (املأ الفراغات *Fill holes*)

الشكل (١-٨) يوضح شبكة الوجه لشخص معين a : اختيار كل النقاط التي يتم اعتيانها بدون الخيار "املأ الفراغات" b: بعد ملأ الفراغات واختيار تخفيض للدقة بمقدار ١/٤



الشكل (١-٨) b :شبكة الوجه ثلاثية الأبعاد عند دقة ١/١ بدون ملأ الفراغات وb :نفس الشبكة عند دقة ١/٤ بعد ملأ الفراغات.

كل شبكة (صورة ملتقطة) مؤلفة من نقاط من سطح الوجه والروابط بين هذه النقاط مشكلة خلايا من أربع نقاط لكل خلية. الاحداثيات (x,y,z) للنقاط ثلاثية الأبعاد(٣D points) مأخوذة باعتبار أن مبدأ الاحداثيات هو الماسح. تم تحويل كل شبكة إلى ملف من نوع VRML بدقة ١/١ بدون معلومات الألوان وبعد عملية ملئ الفراغات. بحذف معلومات الألوان يتم تخفيض حجوم الملفات إلى ٠,٧ Mbytes تقريباً مقارنة مع حجمها الأصلي الذي كان ٢,٧ Mbytes. المرجع [٣]

١-٦-٣ تحليل مرحلة الالتقاط :

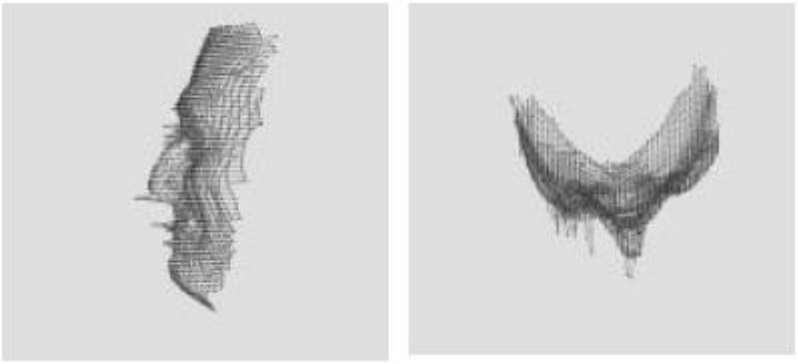
فيما يلي تحليل و مناقشة حول بعض الخواص الهامة للشبكات ثلاثية الأبعاد الملتقطة بالاضافة إلى بعض الخواص المتعلقة بعملية الالتقاط. بعض المصادر المختلفة للخطأ والتشويه في السطوح المولدة هي: وجود الفراغات،اطار وجه غير كامل، وجود الضجيج،...الخ .على الرغم من أنه بالامكان وضع بعض الضوابط خلال مرحلة الالتقاط بهدف انقاص بعض الأخطاء المشار إليها، فإنه لم يتم وضع مثل هذه الضوابط وذلك للحصول على صور أكثر واقعية وعمومية

. وجود الفراغات : الجهاز ثلاثي الأبعاد 3D digitizer هو حساس ليزري فراغي يستخدم تقنية فعالة للحصول على قيم العمق بالاعتماد على اسقاط ضوء متحكم به. وبالتالي فإن الأجزاء المظلمة من السطح لا تعكس الضوء المسقط ولا يمكن اعتيانها مما ينتج عن هذه المناطق فراغات في الشبكة. تابع "ملأ الفراغات" متوفر بواسطة سوفتوير الماسح باستخدام طريقة اضافة interpolation معينة وبالتالي ينتج عن ذلك بعض الخطأ في السطوح المولدة. الشكل (١-٩) يبين في السطر الأول بعض الصور فيها فراغات والسطر الثاني يوضح كيف يتم تخفيض عدد الفراغات عند تطبيق تابع الاضافة "ملأ الفراغات" .

الشكل (١-٩) a: وجود الفراغات في أربع أمثلة.b : نفس الصور بعد مرحلة ملئ الفراغات .

. وجود مناطق مغطاة بالذقن أو الشوارب. هناك بعض الأشخاص في قاعدة البيانات الذين لديهم بعض هذه الملامح في الوجه والتي تغطي السطح الحقيقي للوجه. وهذه الأمور تسبب ضجيج أو فراغات حول هذه المناطق. الشكل (١-٩) يبين مثالاً عن ذلك .

. وجود الضجيج : الصور فيها الكثير من نقاط الضجيج والتي يجب إزالتها بدقة بواسطة إضافة مرحلة معالجة أولية في الأنظمة التي تستخدم هذه الصور. الشكل (١-١٠) يبين الشبكة للشكل ١-٨ من وجهتي نظر(جانبية و علوية على التوالي) مبيناً نقاط الضجيج .



الشكل (١-١٠) الشبكة من وجهتي نظر(جانبية و علوية على التوالي) مبيناً نقاط الضجيج

لشرح بعض الخواص المتعلقة بعملية الالتقاط. أولاً يجب أن نشير إلى الغنى الموجود في التغيرات المختفلة في سطح الوجه لكل شخص ليس فقط بالنسبة إلى الوضع ولكن أيضاً المتعلق بشكل الوجه (تعابير الوجه. بشكل خاص)، من المهم لقاعدة بيانات أن يكون كلا من منظر الضحك أو منظر الملامح العشوائية للوجه واضحتين. الشكل ١-١١ يوضح بعض الأمثلة عن الملامح في الشبكات الملتقطة وفق المنظرين السابقين. ونلاحظ أن الشخص قادر على تغطية وجهه جزئياً بواسطة لسانه (b) أو يده (e).

الشكل (١-١١) أمثلة عن الملامح في الشبكات الملتقطة وفق منظري الضحكة أو ملامح عشوائية للوجه

من الشائع في الصور ذات الملامح وجود دوران في الرأس كما هو ملاحظ في الأمثلة في الشكل ١-١١ (a) و(b). تحدث هذه الظاهرة أيضاً في الصور الأمامية الأخرى. من جهة أخرى أحياناً يكون للأشخاص تعبير ناتج عن فرحهم أو عصبيتهم عند التقاط الصور عديمة التعابير. مثل هذه الصور تم الاحتفاظ بها بدلاً من استبدالها بصور أخرى جديدة. المرجع [٣]

الفصل الثاني

الكشف وتصحيح الوضعية الآلي

للوجوه ثلاثية الأبعاد

٢-١ مقدمة:

يتضمن هذا البحث طريقة آلية بشكل كامل للتعرف الثلاثي الأبعاد على الوجه. ونقدم عدة مقترحات لجعل عملية التعرف التي تستفيد من ICP ذات سرعة عالية نسبياً وقوية تجاه التغيرات في تعابير الوجه. وتضم هذه المقترحات:

١)الكشف الآلي للوجه ثلاثي الأبعاد (تحديد منطقة الوجه) بواسطة تحديد قمة الأنف ومحاولة تحسين دقة هذا التحديد.

٢)التصحيح الآلي لوضعية الوجه الثلاثي الأبعاد بواسطة PCA ثم محاولة تخفيض الأخطاء الناتجة بعد عملية التصحيح.

٣)القوة تجاه تعابير الوجه بتجزيء الوجه أتوماتيكياً إلى مناطق حساسة وغير حساسة تجاه التعابير. وتم حساب معدلات التعرف من أجل الوجوه تحت الاختبار بوجود التعابير أو بدون وجود التعابير.

٤)تسريع عملية التعرف باستخدام تكمية للتمثيل ثلاثي الأبعاد لمناطق الوجه الأقل تأثر بالتعابير وفق دوائر متمركزة حول قمة الأنف واستخدام هذا التمثيل لاستبعاد عدد كبير من الوجوه المرشحة خلال عملية المقارنة أثناء التعرف وبشكل سريع وذلك بهدف تسريع عملية التعرف وزيادة كفاءتها.

هناك عدة تحديات في عملية التعرف على الوجوه منها التنوع العرقي والاختلافات الناتجة عن التعابير أو الجنس أو اختلاف الوضعية للوجه أو بسبب اختلاف الاضاءة أو بسبب وجود المكياج على الوجه. خوارزميات التعرف على الوجوه ثنائية الأبعاد المعتمدة على المظهر كانت من أول ما تم البحث فيه بسبب الانتشار الواسع للكاميرات. أحد الطرق الكلاسيكية في التعرف على الوجوه تستخدم التمثيل باستخدام الوجوه الخاصة جداً والمنفذة من قبل Turk and Pentland والتي تعتمد على تحليل المكونات الأساسية(PCA). التحليل التمييزي الخطي وتحليل المكونات المستقلة وطرق بايسيان كلها طرق تم استخدامها بنجاح من أجل عملية التعرف على الوجوه المعتمدة على المظهر.

قدم Zhao تحليلاً مفصلاً لخوارزميات التعرف على الوجوه ثنائية الأبعاد واستنتج أن الخوارزميات الموجودة في هذا المجال حساسة تجاه تغيرات الاضاءة ووضعية الوجه لذلك يقوم الباحثون حالياً بالبحث في طرائق أخرى للحصول على بيانات الوجه للتغلب على هذه المشاكل. أحد الطرائق الواعدة هي الشكل الثلاثي الأبعاد للوجه.

الوجه الثلاثي الأبعاد هو شعاع ثلاثي الأبعاد _[*xi, yi, zi*] للإحداثيات الثلاثة x,y,z لمجموعة نقاط تمثل الوجه (i=١.....n حيث n عدد النقاط). الوجه ثنائي الأبعاد عبارة عن شعاع ذو خمس أبعاد_[*ui, vi,Ri,Gi,Bi*] حيث u,v احداثيات البكسلات و R,G,B المركبات الزرقاء والخضراء والحمراء لهذه البكسلات. قدم Bowyer تحليلاً لخوارزميات التعرف على الوجه ثلاثية الأبعاد واستنتج أن التعرف على الوجه ثلاثي الأبعاد له إمكانية التغلب على المشاكل التي تعاني منها طرق التعرف ثنائية الأبعاد خاصة أن الشكل الثلاثي الأبعاد للوجه يمكن استخدامه لتصحيح وضعية الوجه.

نقدم في هذا البحث خوارزمية آلية بشكل كامل لتصحيح الوضعية للوجه ثلاثي الأبعاد. الطرق المستخدمة حالياً تقوم بتصحيح الوضعية بواسطة التحديد اليدوي لنقاط علام على الوجه. بينما الطريقة المتبعة في هذا البحث هي التحديد الآلي لقمة الأنف ثم تحسين دقة هذا التحديد وتصحيح الوضعية باستخدام PCA ثم تصحيح الأخطاء الباقية بعد التصحيح و تمثيل الدوائر المتمركزة للمناطق الأقل تأثر بالتعابير من الوجه واستخدامه للاستبعاد السريع لترشيح عدد كبير من الوجوه في قاعدة البيانات عند عملية المقارنة والوجوه الباقية يتم استخدامها في عملية التعرف بواسطة خوارزمية تعرف جديدة قوية تجاه التعابير وهذه القوة تأتي بتجزيء الوجه أتوماتيكياً إلى مناطق حساسة (يتأثر شكلها كثيراً بتعابير الوجه مثل الضحكة أو الابتسامة) وغير حساسة تجاه التعابير (أقل تأثراً بتعابير الوجه) واستخدام المناطق غير الحساسة تجاه التعابير من أجل التعرف.

مراحل العمل :

١- مرحلة المعالجة الأولية وتضم مرحلتين:

أ-التحديد أو الكشف الآلي للوجه ثلاثي الأبعاد (تحديد منطقة الوجه) بواسطة تحديد قمة الأنف (face Automatic ٣D detection)

ب-التصحيح الآلي لوضعية الوجه الثلاثي الأبعادبواسطةPCA(Automatic ٣D pose correction) ثم تصحيح الأخطاء الباقية بعد عملية التصحيح.

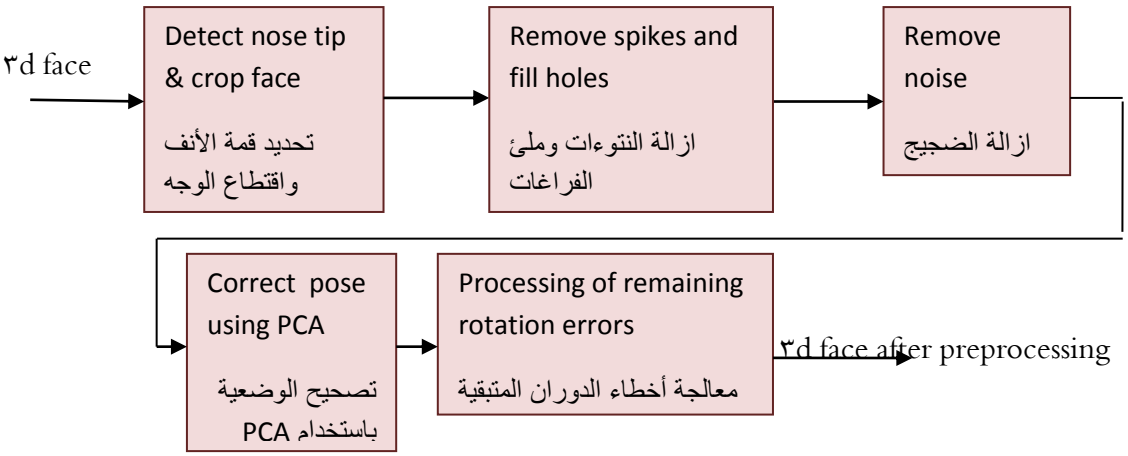
٢- تجزيء الوجه حسب الحساسية تجاه التعابير.

٣- تسريع عمليةالتعرف باستخدام تكمية للتمثيل ثلاثي الأبعاد للمناطق الأقل تأثر بالتعابير من الوجه.

٤-الاستفادة من ICP(iterative closest point) من أجل عملية التعرف(face recognition)

٢-٢- مرحلة المعالجة الأولية(Preprocessing) للوجه ثلاثي الأبعاد:

استخدم Blanz نماذج مرنة للتعامل مع التغيرات في الوضعية في صور الوجوه ثنائية الأبعاد. نستخدم هنا بيانات وجوه ثلاثية الأبعاد GavabDB بهدف التصحيح الآلي لوضعية الوجه. الشكل التالي يبين المخطط الصندوقي لمرحلة المعالجة الأولية التي نقدمها في هذا البحث وسوف نوضح المراحل المختلفة فيما بعد:



الشكل(٢-١) المخطط الصندوقي لمرحلة المعالجة الأولية

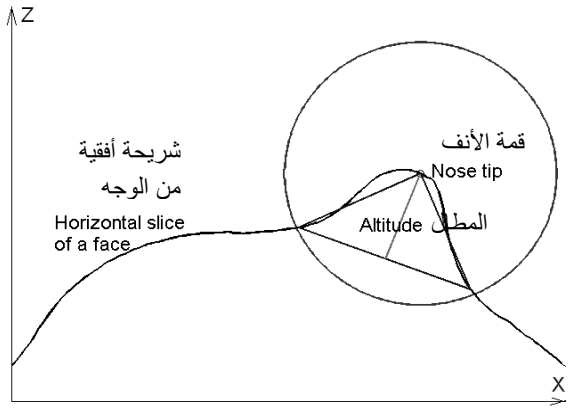
٢-٢-١ تحديد الوجه face detection:

قاعدة البيانات التي نتعامل معها فيها وجوه ثلاثية الأبعاد كلها مأخوذة تقريباً قريباً من مستوى الكتف لذلك من الضروري كخطوة معالجة أولية القيام بتحديد منطقة الوجه.

بما أن معالجة البيانات الثلاثية الأبعاد يحتاج لجهد رياضي كبير فإننا نقوم باقتطاع منطقة الوجه المطلوبة من الوجه الثلاثي الأبعاد من أجل القيام بعمليات المعالجة عليه في المراحل التالية ونقوم بتحديد قمة الأنف في الخطوة الأولى.

طريقة تحديد قمة الأنف: يتم أخذ شرائح أفقية من كل وجه ثلاثي الأبعاد بخطوة dv وفي البداية يتم اختيار قيمة كبيرة ل dv لتحسين سرعة الخوارزمية (و حالما يتم تحديد قمة الأنف بشكل تقريبي تتم اعادة البحث في المنطقة المجاورة لقمة الأنف بقيمة أصغر ل dv).

بعدئذ يتم استخدام دوائر متمركزة عند فواصل أفقية d_h في الشريحة لاختيار قطعة من الشريحة ويتم رسم مثلث باستخدام مركز الدائرة ونقاط تقاطع الشريحة مع الدائرة كما هو موضح بالشكل :



الشكل (٢-٢):تحديد قمة الأنف في شريحة slice

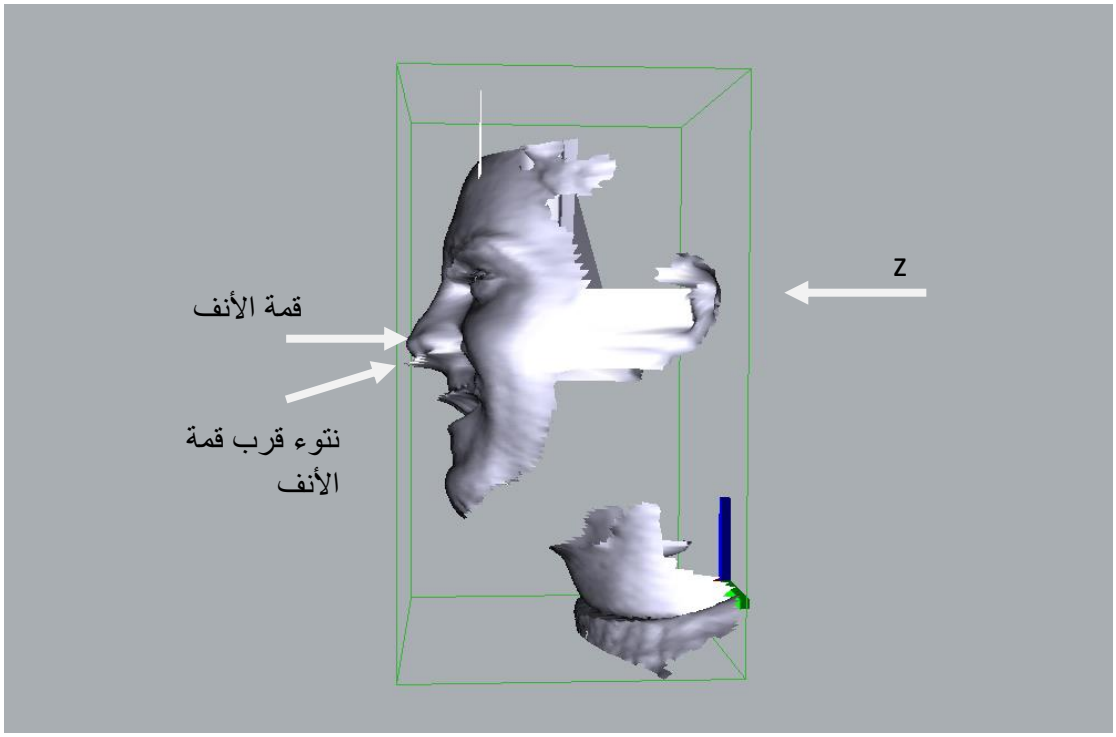
وأيضاً يتم استخدام قيمة كبيرة لd_h في البداية (ثم وبعد التحديد الدقيق لقمة الأنف يتم اختيار قيمة أصغر ل d_h من أجل الدقة في تحديد قمة الأنف).

يتم حساب مطال كل نقطة على أنه البعد بين رأس المثلث المرسوم وقاعدته والنقطة ذات المطال الأكبر تعتبر قمة محتملة للأنف في الشريحة ويتم اسناد قيمة مرجعية لها تساوي المطال.

يتم تكرار هذه العملية من اجل كل الشرائح فنحصل على نقطة مرشحة لتكون قمة الأنف في كل شريحة مع القيمة المرجعية لهذه النقطة. هذه النقاط المرشحة لتكون قمة الأنف تشكل حافة الأنف. يتم المقارنة بين القيم المرجعية لهذه النقاط(مطالات النقاط) والنقطة ذات المطال الأكبر تكون قمة الأنف وتتم اعادة العملية السابقة من أجل فواصل dv و dh أصغر في المنطقة المجاورة لقمة الأنف من أجل التحديد الأدق لقمة الأنف .

٢-٢-١-١ المشاكل التي تواجه هذا الأسلوب لتحديد قمة الأنف:

-وجود ضجيج على المحور z على شكل نتوءات كما يوضح ذلك الشكل التالي:



الشكل (٢-٣) وجه يحوي نتوء ناتج عن الضجيج

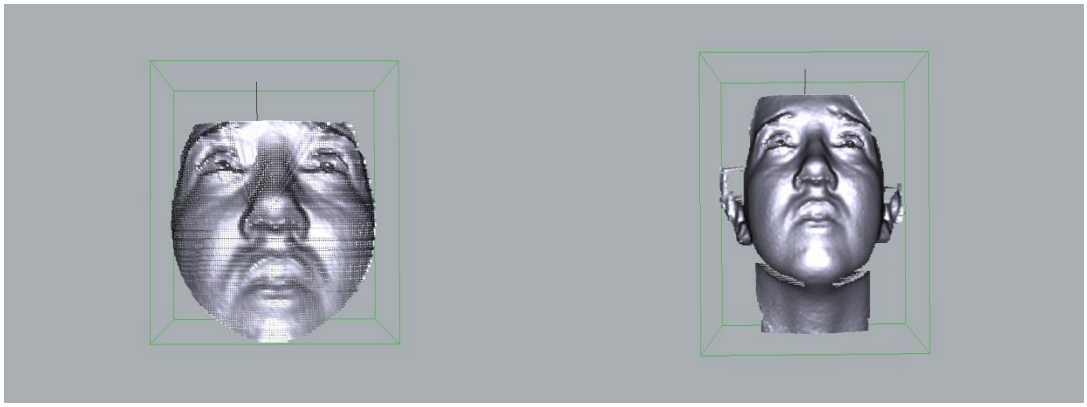
ونتيجة وجود مثل هذا النتوء الموضح في الشكل السابق قد نحصل على تحديد خاطئ لقمة الأنف وذلك حسب قيمة z للنتوء حيث يعطي الأسلوب السابق لتحديد قمة الأنف في هذا المثال الموضح بالشكل (٢-٣) أن قمة الأنف هي البروز الناتج عن الضجيج. لمعالجة مثل هذه المشكلة نحدد النقطة ذات المطال الأكبر والنقطة التي تليها في المطال وفق الأسلوب المذكور أعلاه (أي نحدد النقطة المرشحة لتكون قمة الأنف ونقطة أخرى سنفحص امكانية كونها قمة الأنف).

بالأخذ بعين الاعتبار النقاط المرشحة لتكون قمة الأنف في كل شريحة فإننا نلاحظ من الشكل أعلاه أن فرق المطال تدريجي وقليل نسبياً بين نقطة قمة الأنف الصحيحة ومطال النقاط المرشحة في الشرائح الأفقية المجاورة الأعلى وفق المحور y بينما لا يتحقق ذلك من أجل نتوء الضجيج حيث يحدث في الشريحة الأفقية المارة به زيادة كبيرة نسبياً ومفاجئة لمطال النقطة المرشحة لتكون قمة الأنف بالنسبة لمطال نقاط الشرائح المجاورة الأعلى.

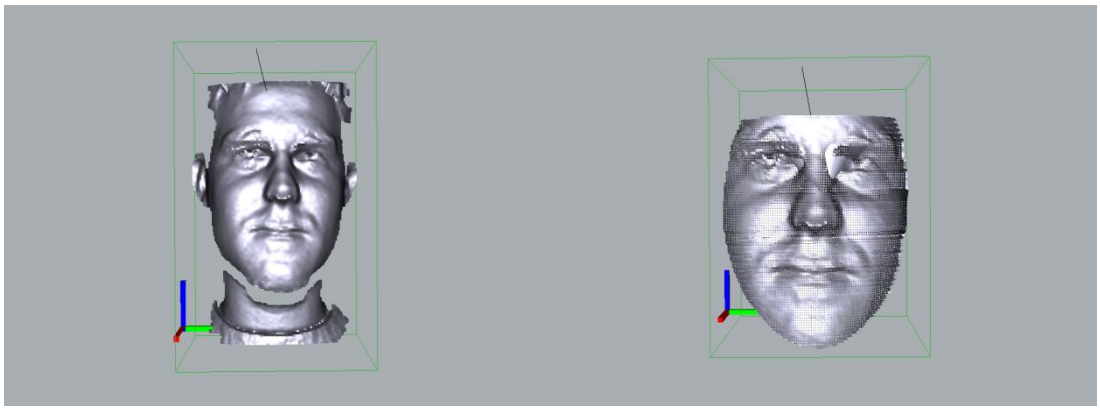
وبالاستفادة من هذه الخاصية نقارن بين النقطتين المرشحتين لتكونا قمة الأنف (النقطتين ذواتي المطال الأكبر) ونحدد قمة الأنف على أنها النقطة ذات الفرق الأقل في المطال عن مطال النقطة المرشحة في الشريحة الأعلى.

٢-٢-١-٢ نتائج مرحلة اقتطاع منطقة الوجه:

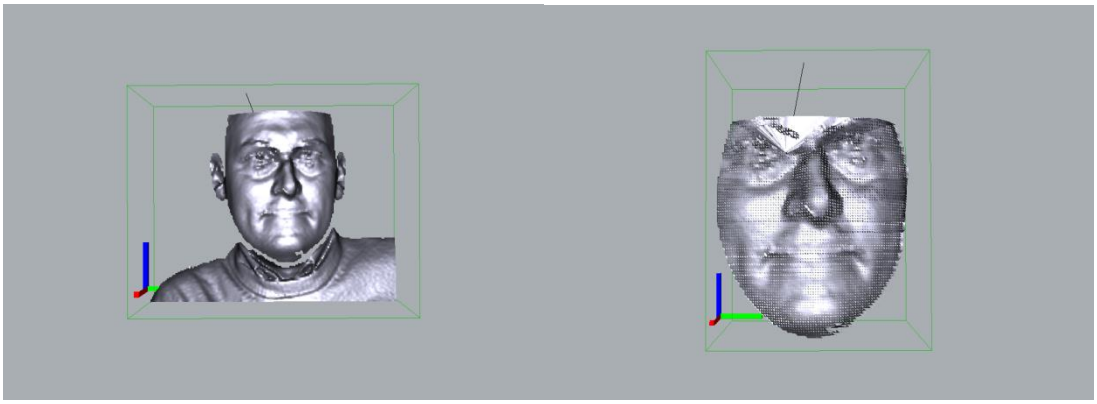
نرسم شكل بيضوي مركزه قمة الأنف لاقتطاع منطقة الوجه الثلاثي الأبعاد وتم اختيار الأبعاد كما يلي: وفق x (٦٠ mm) ووفق y (٧٥ mm) هذه العملية تقتطع منطقة بيضوية من الوجه عند النظر إليها وفق المحورين حيث المحور الأكبر هو المحور الشاقولي y والمحور الأصغر هو المحور الأفقي x كما هو موضح في الأشكال التالية التي توضح تنفيذ هذه الخوارزمية على وجوه ثلاثية الأبعاد في قاعدة بيانات الوجوه ثلاثية الأبعاد لاقتطاع منطقة الوجه :



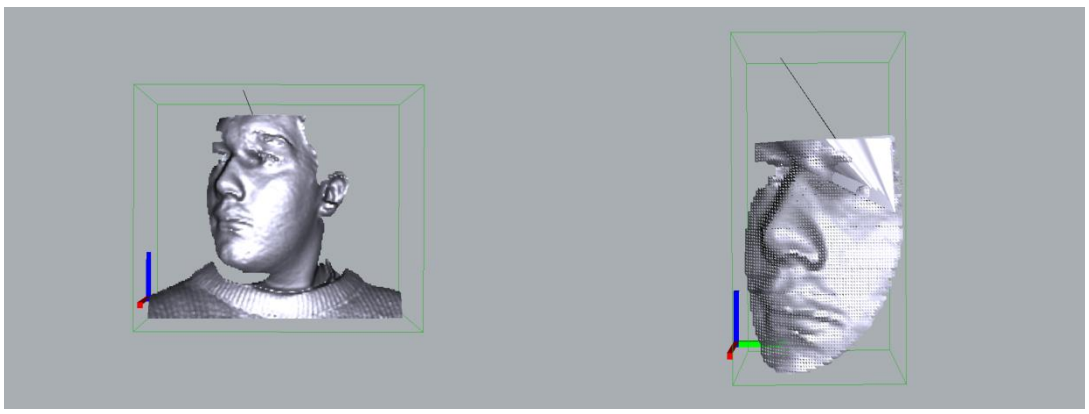
الشكل(٢-٤)الصورة قبل اقتطاع منطقة الوجه **منطقة الوجه بعد الاقتطاع**



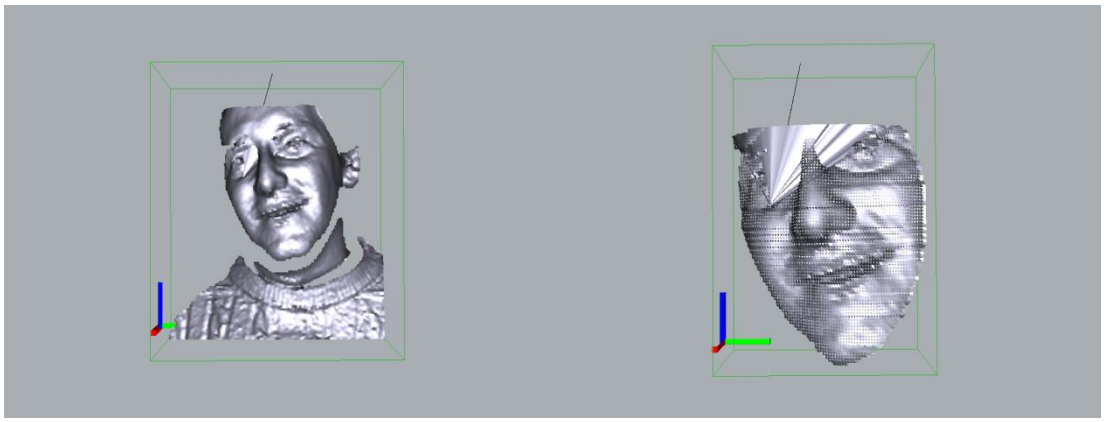
الشكل(٢-٥) منطقة الوجه قبل الاقتطاع منطقة الوجه بعد الاقتطاع



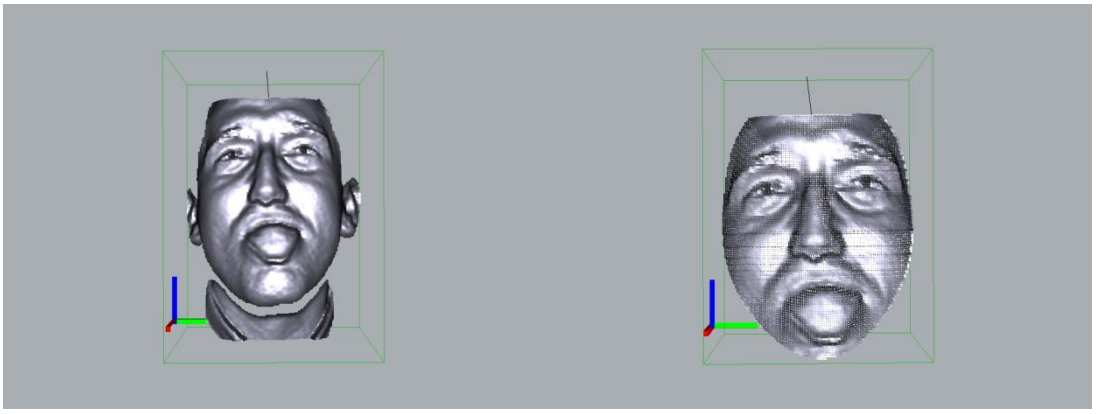
الشكل(٢-٦) منطقة الوجه قبل الاقتطاع منطقة الوجه بعد الاقتطاع



الشكل(٢-٧) منطقة الوجه قبل الاقتطاع منطقة الوجه بعد الاقتطاع



الشكل(٢-٨) منطقة الوجه قبل الاقتطاع **منطقة الوجه بعد الاقتطاع**



الشكل(٢-٩) منطقة الوجه قبل الاقتطاع **منطقة الوجه بعد الاقتطاع**

ونلاحظ قوة هذه الخوارزمية رغم التغيرات في وضعية الوجه والتغيرات في التعابير والتي تظهر في الأشكال السابقة حيث يكون معدل نجاح تحديد قمة الأنف ١٠٠% من أجل الوجوه بوضعية أمامية بتعابير وبدون تعابير والموجودة في قاعدة بيانات الوجوه ثلاثية الأبعاد.

تغيرات الوضعية pose التي تعمل من أجلها طريقة تحديد قمة الأنف تقع في حدود زوايا الدوران التالية : (٤٥ درجة حول المحور y)- (٣٥ درجة حول المحور x)

وهي زوايا الدوران التي من أجلها تنجح مرحلة تصحيح الوضعية pose correction

يتناقص أداء هذه الخوارزمية عندما تصبح زاوية الدوران حول المحور y قريبة من ٤٠ درجة وأكبر حيث يختفي الجانب الآخر من الأنف مما يحول دون امكانية تحديد قمته.

الجدولان التاليان يبينان على سبيل التوضيح النتائج من أجل ٢٠ وجه بوضعيات مختلفة بدون تعابير مع زاوية الدوران حول المحور المذكور ثم نضع بعد كل جدول معدل التحديد الصحيح لقمة الأنف من أجل قاعدة البيانات المؤلفة من ٣٢ وجه حيث :

success:تشير إلى عملية تحديد ناجحة لقمة الأنف ،failure:تشير إلى عملية تحديد فاشلة لقمة الأنف

أولاً : الدوران حول المحور y :

رقم الشخصية	وضعية١(دوران أقل من ٤٠ درجة نحو اليسار)	وضعية٢ (دوران أقل من ٤٠ درجة نحو اليمين)	وضعية جانبية دوران نحو اليسار	وضعية جانبية دوران نحو اليمين
١	success(٢٨درجة)	success(٢٧درجة)	failure	success
٢	success(٢٩درجة)	success(٢٤درجة)	failure	failure
٣	success(٢٧درجة)	success(٢٠درجة)	failure	failure
٤	success(٢٦درجة)	success(٢٥درجة)	failure	failure
٥	success(٣١درجة)	success(١٩درجة)	failure	failure
٦	success(٢٥درجة)	success(٢٧درجة)	failure	failure
٧	success(٢٢درجة)	success(٢٦درجة)	failure	failure
٨	success(٢٠درجة)	success(١٨درجة)	failure	failure
٩	success(١٨درجة)	success(١٧درجة)	failure	failure
١٠	success(١٦درجة)	success(١٨درجة)	success	failure
١١	success(١٧درجة)	success(٢٦درجة)	failure	failure
١٢	success(٢٧درجة)	success(٢٣درجة)	failure	failure
١٣	success(٢٤درجة)	success(١٨درجة)	success	failure
١٤	success(٢١درجة)	success(٣٣درجة)	failure	failure
١٥	success(٢٠درجة)	success(٢٠درجة)	failure	failure
١٦	success(٢٥درجة)	success(١٨درجة)	failure	failure
١٧	success(٢٣درجة)	success(٢١درجة)	failure	failure
١٨	success(٢٧درجة)	success(١٩درجة)	success	failure
١٩	success(٢٥درجة)	success(١٤درجة)	failure	success
٢٠	success(١٩درجة)	success(٢٢درجة)	failure	failure

الجدول(٢-١)

معدل التحديد الصحيح لقمة الأنف من أجل الوجوه بوضعيات مختلفة (دوران حول المحور y أقل من ٤٥ درجة) (١٠٠%)

معدل التحديد الصحيح لقمة الأنف من أجل الوجوه بوضعيات مختلفة (دوران حول المحور y أكبر من ٤٥ درجة أي وضعيات جانبية profile للوجه) (١٧,١٩%)

ثانياً : الدوران حول المحور x :

رقم الشخصية	وضعية جانبية١ (دوران حول المحور xنحو الأعلى)	وضعية جانبية٢(دوران حول المحور xنحو الأسفل)
١	success(٢٠درجة)	success(٢٠درجة)
٢	success(١٥درجة)	success(١٨درجة)
٣	success(١٥درجة)	success(٤درجة)
٤	success(٣٠درجة)	success(٢٠درجة)
٥	success(٢٣درجة)	success(١٦درجة)
٦	success(٢٢درجة)	success(٣درجة)
٧	success(١٩درجة)	success(٣٠درجة)
٨	success(٢٤درجة)	success(٢٤درجة)
٩	success(١١درجة)	success(٦درجة)
١٠	success(١٩درجة)	success(٤درجة)
١١	success(٢٨درجة)	success(٢١درجة)
١٢	success(١١درجة)	success(١٧درجة)
١٣	success(٢١درجة)	success(٩درجة)
١٤	success(٢٢درجة)	failure(بسبب الشعر على الوجه)
١٥	success(١٧درجة)	success(١٨درجة)
١٦	success(٣٢درجة)	success(١٣درجة)
١٧	failure	success(١٧درجة)
١٨	failure	success(٢٠درجة)
١٩	failure	success(٢٣درجة)
٢٠	success(٢٥درجة)	success(٢٥درجة)

الجدول(٢-٢)

معدل التحديد الصحيح لقمة الأنف من أجل الوجوه بوضعيات مختلفة (دوران حول المحورx حتى ٣٥ درجة) (٩٠,٦٣%)

دقة عملية تحديد قمة الأنف من أجل وجوه بوضعية أمامية ووضعيات دوران حول المحور y حتى ٤٥ درجة وحول المحور x حتى ٣٥ درجة هي ٩٦,٨٨% (٦ حالات فشل من أصل ١٩٢ وجه). والأخطاء كانت ناتجة إما عن الشعر الذي يغطي جزء من الوجه أو اقتراب زاوية الدوران حول المحور x من ٣٥ درجة أو بسبب النتوءات الكبيرة والتي تعطي قيم z كبيرة غير صحيحة.

وسبب انخفاض معدل التحديد الصحيح لقمة الأنف من أجل زوايا دوران حول المحور y أكبر من ٤٥ درجة هو بسبب زيادة زوايا الدوران إلى حد يغطي الجانب الآخر من الأنف (وضعيات قريبة من الوضعية الجانبية للوجه profile) مما يحول دون امكانية تحديد قمته حيث تفشل خوارزمية تحديد قمة الأنف المشار إليها أعلاه حيث نحصل على نقطة تقاطع واحدة بين الدائرة والشريحة من أجل النقاط القريبة من حافة الأنف وقمة الأنف مما يحول دون امكانية تحديد مطال تلك النقاط.

حالما يتم اقتطاع الوجه يتم ازالة النقاط التي تشكل نتوءات غير صحيحة في الوجه ثلاثي الأبعاد(spikes) و هذه النقاط الخارجية هي النقاط التي يكون بعدها عن أي نقطة من النقاط المجاورة أكبر من عتبة dt. يتم حساب dt أتوماتيكياً باستخدام $dt = \mu + ٠.٦\sigma$ حيث (μ هي المسافة المتوسطة بين النقاط المتجاورة و σ هي الانحراف المعياري). ينتج عن ازالة هذه النتوءات فراغات في الوجه ثلاثي الأبعاد يتم ملؤها بواسطة متوسط قيم النقاط المجاورة.

إلا أن هذه الخطوة تستهلك وقتاً كبيراً بسبب العدد الكبير للنقاط في السطح ثلاثي الأبعاد لذلك يمكن الاستغناء عن هذه الخطوة طالما كانت هذه النتوءات ضمن الحد المقبول وفي هذا البحث لم تؤثر هذه النتوءات على نتائج عملية التعرف مما أمكننا من تخطي هذه الخطوة.

بما أن الضجيج في البيانات ثلاثية الأبعاد يحدث عادة على طول المحور z بالنسبة للحساس لذلك نزيل الضجيج من المكون z للوجه ثلاثي الأبعاد باستخدام المرشح (median filter).

٢-٢-٢ تصحيح الوضعية :pose correction

٢-٢-٢-١ المحاور الأساسية:

مع التعريف الأكثر وضوحاً لهدف PCA نحتاج أيضاً لتعريف أكثر دقة للبيانات التي بين أيدينا. نقوم بمعاملة كل عينة في الزمن كعينة مستقلة في مجموعة البيانات لدينا حيث في كل لحظة اعتيان نسجل مجموعة من البيانات تتألف من مجموعة من القياسات (مثلاً الجهد،الموضع،الخ). فلو فرضنا مثال حركة نابض مسجلة بواسطة ثلاث كاميرات فإنه وفي كل لحظة اعتيان تسجل الكاميرا الأولى موضع الكرة المربوطة بالنابض(xa,ya) وبالتالي يمكن التعبير عن كل عينة بشعاع ذو ٦ أبعاد.

$$\vec{X} = \begin{bmatrix} x_A \\ y_A \\ x_B \\ y_B \\ x_C \\ y_C \end{bmatrix}$$

حيث كل كاميرا تعطي مسقطين لموضع الكرة من الشعاع الإجمالي X وإذا سجلنا موضع الكرة من أجل ١٠ دقائق عند تردد ١٢٠Hz فسيكون لدينا ١٠*٦٠*١٢٠=٧٢٠٠٠ من هذه الأشعة.

كل عينة X عبارة عن شعاع ذو m بعد حيث m عدد أنواع القياس. أي أن كل عينة هي شعاع في فضاء أشعة ذو m بعد له أشعة واحدية متعامدة m ونعلم من الجبر الخطي أن كل أشعة القياس عبارة عن جمع خطي من هذه المجموعة من أشعة الواحدة فما هي هذه الأشعة الواحدية المتعامدة؟

لنفترض أننا جمعنا البيانات الخاصة بمثال النابض لكننا نظرنا فقط إلى الكاميرا A فما هي أشعة الواحدة المتعامدة من أجل (xa,ya)؟ سيكون الجواب البديهي {(١،٠),(٠،١)} لكن لماذا لا نختار أشعة الواحدة{(aqrt(٢)/٢,sqrt(٢)/٢),(-sqrt(٢)/٢, sqrt(٢)/٢)} أو لماذا لا نختار أي دوران آخر؟ السبب هو أن المحاور الأساسية البديهية تعكس طريقة حصولنا على البيانات. فلو فرضنا أننا سجلنا الموضع (٢،٢) فإننا لا نعبر عن الموضع بـ٢*sqrt(٢) في الاتجاه (aqrt(٢)/٢,sqrt(٢)/٢) و٠ بالاتجاه المعامد وبدلاً عن ذلك نسجل الموضع(٢،٢) بواسطة كاميرتنا مما يعني وحدتين إلى الأعلى ووحدتين إلى اليمين وهكذا فإن المحاور الأساسية البديهية تعكس طريقة حصولنا على البيانات.

كيف نعبر عن هذه المحاور الأساسية البديهية في الجبر الخطي؟ في حالة الفضاء ثنائي الأبعاد يمكن التعبير عن {(١،٠),(٠،١)} بأشعة أسطر منفصلة والمصفوفة المشكلة من شعاعي السطر السابقين هي المصفوفة الواحدية I بالأبعاد ٢*٢. نستطيع تعميم هذا الأمر على فضاء ذو m بعد بتشكيل مصفوفة واحدية m*m :

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I \qquad (١-٢)$$

حيث كل سطر عبارة عن شعاع واحدي bi ذو m مكون ومتعامد مع الأشعة الواحدية(الأسطر) الباقية. نعتبر المحاور البديهية الأساسية على أنها نقطة الانطلاق الفعالة وكل البيانات التي بين أيدينا تم الحصول عليها وفق هذه المحاور وبالتالي يمكن التعبير عنها كجمع خطي من الأشعة bi. المرجع[١]

٢-٢-٢-٢ تغيير المحاور باستخدام تحليل المكونات الأساسية

:(PCA)Principal Component Analyses

تقوم PCA بالبحث عن محاور أخرى والتي تكون عبارة عن جمع خطي من المحاور الأصلية وهذه المحاور الجديدة تعيد تمثيل البيانات بشكل أفضل.

"الخطية" تبسط المسألة عن طريق :١-تحديد مجموعة المحاور الممكن اعتمادها ٢-افتراض الاستمرارية في مجموعة البيانات.

مع هذا الافتراض تكون مهمة PCA محصورة في اعادة تمثيل البيانات والتي تكون جمع خطي للأشعة الواحدية للمحاور. لتكن X مجموعة البيانات الأصلية حيث كل عمود عبارة عن عينة وحيدة من البيانات وأبعاد هذه المصفوفة m*n ولتكن Y مصفوفة أخرى m*n تنتج عن X بواسطة تحويل خطي P أي أن X تمثل مجموعة البيانات الملتقطة الأصلية وتمثل Y تمثيل جديد لمجموعة البيانات.

$$PX = Y \qquad (٢-٢)$$

لنعرف المقادير التالية :

- Pi هي أسطر المصفوفة P

- xi هي أعمدة المصفوفة X

- yi هي أعمدة المصفوفة Y

العلاقة السابقة(٢-٢) تمثل تغييراً للمحاور وبالتالي يمكن تفسيرها بأكثر من أسلوب:

١- P هي مصفوفة تحول X إلى Y

٢- هندسياً،P تمثل مصفوفة دوران وامتداد stretch والتي تحول X إلى Y

٣- أسطر المصفوفة P :{p١,p٢,.....,pm} هي مجموعة الأشعة الواحدية للمحاور الجديدة التي ستمثل أعمدة المصفوفة X (أي التي ستعيد تمثيل مجموعة البيانات X).

التفسير الأخير غير واضح لكن يمكن توضيحه بكتابة :

$$PX = \begin{bmatrix} p_1 \\ \vdots \\ p_m \end{bmatrix} \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$$

$$Y = \begin{bmatrix} p_1 \cdot x_1 & \cdots & p_1 \cdot x_n \\ \vdots & \ddots & \vdots \\ p_m \cdot x_1 & \cdots & p_m \cdot x_n \end{bmatrix} \qquad (٣-٢)$$

يمكن ملاحظة أن كل عمود من المصفوفة Y له الشكل:

$$y_i = \begin{bmatrix} p_1 \cdot x_i \\ \vdots \\ p_m \cdot x_i \end{bmatrix} \qquad (٤-٢)$$

يمكن ملاحظة أن كل معامل من yi عبارة عن ضرب xi مع السطر المقابل في المصفوفة P أي بكلمات أخرى المعامل j من y هو المسقط على السطر j (المحور j) من المصفوفة P. وهذا في الواقع يعطي ان yi هو المسقط على المحاور $\{p_1,\ldots,p_m\}$ لذلك فإن أسطر P في الواقع هي مجموعة جديدة من الأشعة الواحدية لتمثيل أعمدة X

بافتراض الخطية فإن المسألة تصبح ايجاد التغيير المناسب للمحاور. أشعة الأسطر $\{p_1,\ldots,p_m\}$ في تحويل PCA ستصبح المكونات الأساسية لX. وتطرح الآن عدة نقاط:

- ما هي أفضل طريقة لإعادة تمثيل X

- ما هو الاختيار الجيد للمصفوفة P

يجب الاجابة على الأسئلة عن طريق معرفة ما هي الخواص التي يجب أن تكون لمجموعة البيانات وفق التمثيل الجديد Y. وسنحدد هذه الميزات المطلوبة في الفقرة التالية.

التبعثر variance:

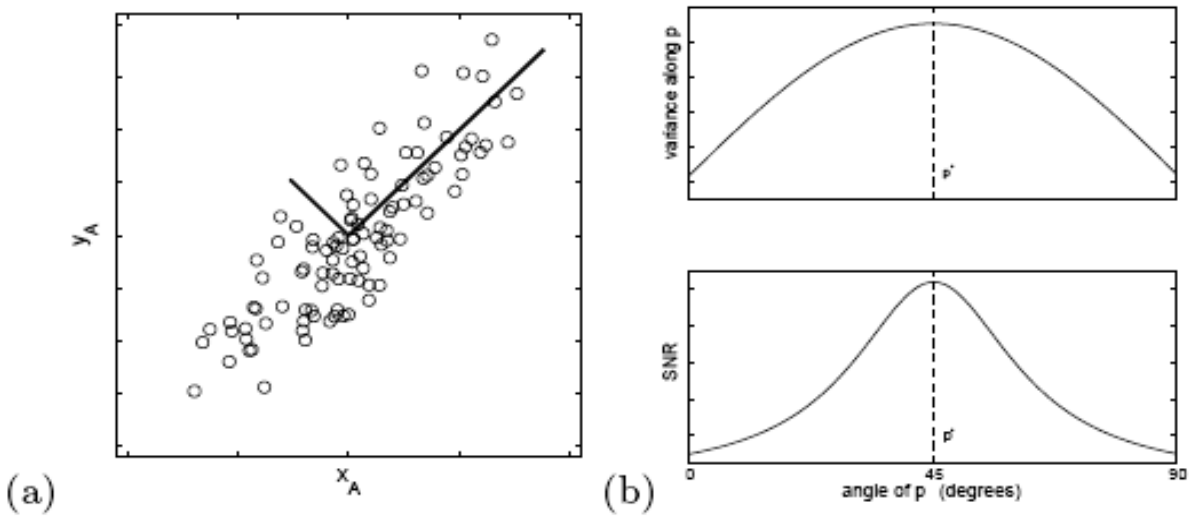
يجب البحث عن ما يمثل متوسط البيانات بشكل أفضل وهذه الفقرة ستتوصل إلى الأسلوب الرياضي اللازم لإعادة تمثيل بيانات غير مثالية في تمثيلها.

أولاً-الضجيج والدوران: Rotation and Noise

الضجيج المقاس في أية مجموعة بيانات يجب أن يكون منخفضاً وإلا لن يكون ممكناً الحصول على معلومات عن النظام. لا يوجد مقياس مطلق للضجيج لكن يتم قياس كل الضجيج بالنسبة للكميات المقاسة من النظام حيث تستخدم نسبة الاشارة إلى الضجيج (SNR) أو نسبة التباين(التبعثر) variance(σ^2 مربع الانحراف المعياري) :

$$SNR = \frac{\sigma^2_{signal}}{\sigma^2_{noise}} \qquad (٥-٢)$$

نسبة SNR الكبيرة (SNR>>١) تعني أنه لدينا بيانات دقيقة بينما نسبة منخفضة لSNR تعني بيانات فيها ضجيج.



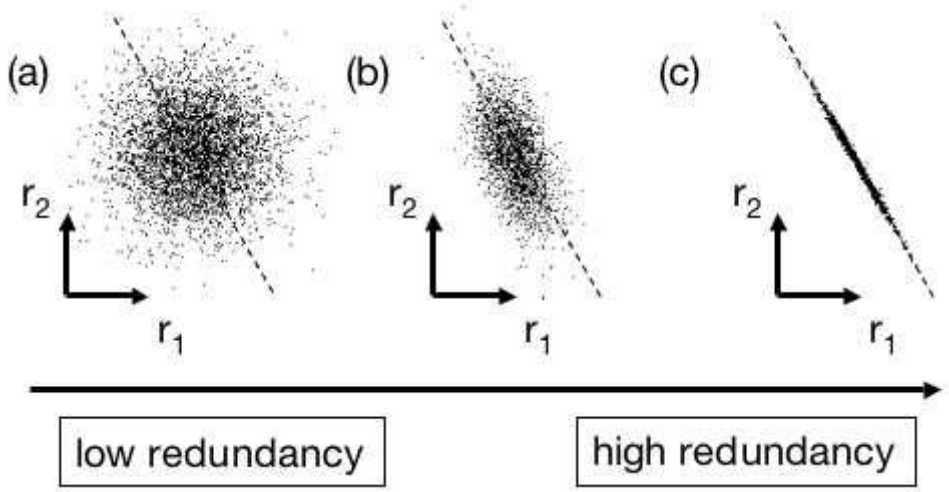
الشكل(٢-٣) فيه: a: بيانات ذات احداثيات (xa,ya). التبعثر للاشارة والضجيج σ^2_{signal} و σ^2_{noise} موضحان بالخطين الممتدين عبر غيمة البيانات، b: تدوير هذه المحاور لإيجاد*p المثالية عندما يصبح التبعثر وSNR أعظميتين. تعرفSNR على أنها نسبة التبعثر وفق*p إلى التبعثر وفق الاتجاه المتعامد.

لنفترض أننا رسمنا كل البيانات لحركة الكرة المعلقة بالنابض وفق الاحداثيات (xa,ya) ولنتذكر أن النابض يسير وفق طريق مستقيم فإن الكاميرا يجب أن تسجل الحركة وفق طريق مستقيم. لذلك فإن أي انحراف للنقاط خارج المسار المستقيم يجب أن يكون ضجيج. التبعثرvariance الناتج عن الاشارة والضجيج يعبر عنهما بواسطة الخطين في الشكل أعلاه وبالتالي النسبة بين طول الخطين (SNR) تقيس كم تكون الغيمة سمينة حيث لدينا عدة حالات منها (SNR>>١) حيث نحصل على غيمة نقاط على شكل خط رفيع أو حالة SNR=١ حيث نحصل على غيمة نقاط دائرية تماماً. باعتبار أنه لدينا قياسات جيدة فإننا نفترض كمياً أن الاتجاهات ذات التبعثر variance الأكبر في فضاء القياس تحوي التغيرات المهمة. في الشكل السابق يكون الاتجاه ذو التبعثر variance الأكبر ليس (١،٠)=xa وليس (٠،١)=ya ولكن يكون وفق اتجاه المحور الأطول لغيمة النقاط. وبالتالي نفترض أن التغيرات المهمة تكون على الاتجاهات ذات التبعثر variance الأكبر أي عند SNR العظمى.

المحاور التي نبحث عنها ليست المحاور الأساسية البديهية (xa,ya) لأن هذه الاتجاهات لا تمثل الاتجاهات ذات التبعثر variance الأكبر. لزيادة التبعثر variance وكذلك نسبة SNR يجب ايجاد الدوران المناسب للمحاور الأساسية أي يجب ايجاد الاتجاه *p في الشكل b في الشكل السابق. في الحالة ثنائية الأبعاد في الشكل a في الشكل السابق تكون *p في اتجاه خط أفضل ملاءمة لغيمة النقاط وبالتالي فإن تدوير المحاور الأساسية لتصبح موازية ل*p سيظهر اتجاه حركة النابض من أجل الحالة ثنائية الأبعاد. يجب تعميم هذا المبدأ على عدد عشوائي من الأبعاد وقبل معرفة الطريقة لذلك سنستعرض هذه المسألة من وجهة نظر أخرى.

ثانياً – التمثيل الزائد:Redundancy

الشكل السابق يبين مشكلة أخرى في تمثيل البيانات فهل من الضروري فعلاً تسجيل متغيرين. الشكل التالي يبين منحنيات ممكنة بين متغيرين مختلفين r_1, r_2.



الشكل (٢-٤) يبين حالات التمثيل الزائد الممكنة من المتغيرين r_1 و r_2. الخط الأكثر ملاءمة $r_2=kr_1$ موضح بالخط المتقطع.

الشكل a: يوضح متغيرين بدون علاقة واضحة بينهما أي بكلمات أخرى r١ غير مترابطة مطلقاً بـr٢ ولأنه لا يمكن توقع r١ من r٢ فإننا نقول أن r١ و r٢ مستقلة خطياً.

من جهة أخرى الشكل c يوضح متغيرين مترابطين بشكل كبير. ومن الواضح أنه من الممكن فقط الاقتصار على متغير واحد وليس متغيرين لأنه من الممكن حساب r١ من r٢ وبالعكس باستخدام خط الملاءمة الأفضل. وهذه هي الفكرة وراء "تخفيض الأبعاد".

ثالثاً – مصفوفة التغاير covariance matrix :

من السهل كشف حالة التمثيل الزائد في حالة وجود متغيرين بواسطة ايجاد ميل خط الملاءمة الأفضل والحكم على جودة هذه الملاءمة. كيف نعمم هذا المبدأ من أجل حالة ذات أبعاد أكبر من ٢؟ لتكن لدينا مجموعتي القيم لمتغيرين ذواتي قيمة متوسطة صفرية:

$$A=\{a_1,a_2,\ldots\ldots,a_n\},B=\{b_1,b_2,\ldots\ldots,b_n\}$$

التبعثر variance لكل من A و B يعرفان كما يلي :

$$\sigma^{\wedge 2}_A=\langle a_i a_i\rangle_i \ , \ \sigma^{\wedge 2}_B=\langle b_i b_i\rangle_i \quad (٦-٢)$$

حيث الرمز <> يشير إلى القيمة المتوقعة أي متوسط الn قيمة للمتغيرات. التغاير covariance بين A و B يعطى بالعلاقة :

$$\text{Covariance of A and B}=\sigma^2_{AB}=\langle a_i b_i\rangle_i \quad (٧-٢)$$

يقيس التغاير covariance درجة العلاقة الخطية أو الترابط بين المتغيرين فعندما تكون قيمتها كبيرة فهذا يعني وجود تمثيل زائد كبير وبالعكس عندما تكون قيمتها صغيرة فهذا يعني وجود تمثيل زائد قليل. في حالة الشكلين b و c في الشكل السابق تكون قيم التغاير كبيرة .

بعض خصائص التغاير covariance :

١) σ^2_{AB} تكون صفراً إذا وفقط إذا كانت A و B غير مترابطتين مطلقاً.

٢) $\sigma^2_{AB}=\sigma^2_A \text{ if } A=B$

يمكن تمثيل A و B على شكل أشعة سطر:

$a = [a_1\ a_2\ \dots\ a_n]$

$b = [b_1\ b_2\ \dots b_n]$

يمكن بالتالي التعبير عن التغاير covariance بعملية ضرب مصفوفية متقابلة :

$$\sigma^2_{\mathbf{ab}} \equiv \frac{1}{n-1}\mathbf{a}\mathbf{b}^T \qquad (٨-٢)$$

حيث $1/(n-1)$ ثابت من أجل التقييس normalization

أخيراً، يمكن التعميم من حالة البعدين (شعاعين) إلى عدد ما من الأبعاد. يمكن اعادة تسمية أشعة السطر $x_1=a$ و $x_2=b$ ولنعتبر أنه لدينا أشعة أسطر اضافية $x_3,\dots,x_m$. يمكن الآن تعريف مصفوفة جيدة X أبعادها m*n:

$$\mathbf{X} = \begin{bmatrix} \mathbf{x_1} \\ \vdots \\ \mathbf{x_m} \end{bmatrix}$$

أحد تفسيرات X ما يلي :كل سطر في X يمثل كل القيم لمتغير محدد (xi) وكل عمود في X يمثل قيمة واحدة لأنواع مختلفة من المتغيرات.نصل الآن إلى تعريف مصفوفةالتغايرcovariance matrix (Cx):

$$\mathbf{C_X} \equiv \frac{1}{n-1}\mathbf{X}\mathbf{X}^T \qquad (٩-٢)$$

العنصر ij من المصفوفة Cx عبارة عن الجداء المتقابل بين شعاع المتغير i وشعاع المتغير j.

بعض خواص المصفوفة Cx :

- Cx مصفوفة مربعة أبعادها m*m .

- العناصر القطرية ل Cx هي التبعثر variance للمتغيرات.

- العناصر غير القطرية لCx هي التغاير covariance بين متغيرات مختلفة.

Cx تظهر الترابط بين كل أزواج المتغيرات وقيم الترابط هذه تعكس التمثيل الزائد redundancy في مجموعة البيانات.

- القيم الكبيرة للعناصر القطرية في المصفوفة تقابل الاشارة المهمة والقيم الصغيرة تقابل الضجيج.

- القيم الكبيرة في العناصر غير القطرية تقابل تمثيل زائدredundancy كبير(ترابط كبير) والقيم الصغيرة تقابل تمثيل زائد قليل(ترابط قليل). المرجع [١]

جعل مصفوفة التغاير مصفوفة قطرية :

يمكن أن نلخص ما سبق بأن هدف PCA هو:

١-تقليل التمثيل الزائد redundancy (تقليل الترابط بين الاحداثيات) والذي يقاس بالتغاير covariance

٢- زيادة الاشارة والمقاسة بالتبعثر variance

التغاير موجب دوماً وبالتالي فإن القيمة الصغرى للتغاير هي الصفر وبالتالي فإن الشكل المثالي لمصفوفة التغاير covariance matrix عندما تكون القيم غير القطرية صفرية أي أن C_Y يجب أن تكون قطرية. لتكن X مجموعة من k شعاع ذو بعد n : $X_1, X_2, \ldots\ldots, X_k$

الشعاع المتوسط لX :

$$m_x = \frac{1}{K}\sum_{i=1}^{K} X_i \qquad (٢-١٠)$$

ومصفوفة التغاير covariance matrix للشعاع الاجمالي X :

$$\begin{aligned} C_x &= E[(X - m_x)(X - m_x)^T] = \\ &= E(XX^T) - E(Xm_x^T) - E(m_xX^T) + E(m_xm_x^T) = \\ &= E(XX^T) - m_xm_x^T - m_xm_x^T + m_xm_x^T = \\ &= \frac{1}{K}\sum_{i=1}^{K} X_iX_i^T - m_xm_x^T \end{aligned}$$

C_x مصفوفة حقيقية متناظرة n*n(حقيقية لأنها مجموع جداءات الأشعة مع منقولها وعناصر الأشعة عبارة عن أعداد حقيقة) لذلك لهذه المصفوفة n قيمة خاصة حقيقة (n real eigenvalues) و n شعاع خاص حقيقي متعامد(n real orthogonal eigenvectors)

لتكن $\lambda_1, \lambda_2, \ldots, \lambda_n$ ال n قيمة خاصة حيث : $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ و $V_1, V_2, \ldots, V_n$ المجموعة المقابلة من الأشعة الخاصة المتعامدة حيث V_1 تقابل λ_1(القيمة الخاصة الأكبر).

نعرف المصفوفة A التي صفوفها هي ال n شعاع خاص $V_1, V_2, \ldots, V_n$:

$$A = \begin{bmatrix} V_1 - - - - > \\ V_2 - - - - > \\ \cdot \\ \cdot \\ V_n - - - - > \end{bmatrix}$$

PCA هو التحويل الخطي الذي يعطى بالعلاقة التالية :

$$Y_i = A(X_i - m_x) \qquad i = 1, 2, \ldots, k \qquad (٢-١١)$$

الشعاع المتوسط للشعاع الاجمالي الناتج Y هو ٠ بما أن :

$$\begin{aligned} m_y &= \frac{1}{K}\sum^{K} Y_i = \\ &= \frac{1}{K}\sum_{i=1}^{K} A(X_i - m_x) = \\ &= A\frac{1}{K}\sum_{i=1}^{K}(X_i - m_x) = 0 \end{aligned}$$

ومصفوفة التغاير للشعاع الاجمالي Y مصفوفة قطرية تعطى كما يلي :

$$\begin{aligned} C_y &= E[(Y - m_y)(Y - m_y)^T] = \\ &= \frac{1}{K}\sum_{i=1}^{K} Y_i Y_i^T - m_y m_y^T = \\ &= \frac{1}{K}\sum_{i=1}^{K} Y_i Y_i^T = \\ &= \frac{1}{K}\sum_{i=1}^{K}[A(X_i - m_x)][A(X_i - m_x)]^T = \\ &= \frac{1}{K}\sum_{i=1}^{K} A(X_i - m_x)(X_i - m_x)^T A^T = \\ &= A[\frac{1}{K}\sum_{i=1}^{K}(X_i - m_x)(X_i - m_x)^T]A^T = \\ &= AC_x A^T = \end{aligned}$$

$$= \begin{bmatrix} V_1 \longrightarrow \\ V_2 \longrightarrow \\ \cdot \\ \cdot \\ V_n \longrightarrow \end{bmatrix} C_x \begin{bmatrix} V_1 & V_2 & \cdot\cdot & V_n \\ \downarrow & \downarrow & \downarrow & \downarrow \\ & & & \end{bmatrix} =$$

$$= \begin{bmatrix} V_1 \longrightarrow \\ V_2 \longrightarrow \\ \cdot \\ V_n \longrightarrow \end{bmatrix} \begin{bmatrix} \lambda_1 V_1 & \lambda_2 V_2 & \cdot\cdot & \lambda_n V_n \\ \downarrow & \downarrow & \downarrow & \downarrow \\ & & & \end{bmatrix} =$$

$$= \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \cdot & \\ 0 & & & \lambda_n \end{bmatrix}$$

حيث $V_i \longrightarrow$ يمثل شعاع صف و $\downarrow V_i$ يمثل شعاع عمود والمساواة الأخيرة تنتج عن أن الأشعة الخاصة متعامدة وذات طول واحدي .

٢-٢-٢-٣ التحليل إلى المكونات الأساسية PCA والمحاور الأساسية لصورة :

(principal component analysis) (PCA) عبارة عن تحويل خطي لمجموعة من أشعة من بعد n والتي نحصل بنتيجته على n احداثيات غير مترابطة uncorrelated. عند تطبيق PCA على صورة ذات بعدين فإن الصورة المحولة سيتم ربطها مع محاورها الأساسية.

يمكن استخدام PCA لتحسين القدرة على معالجة التشويه في الصورة الناتج عن الدوران والانزياح ولكن هذا التحسن في المقابل يزيد في الوقت اللازم للحصول على الصورة ومعالجتها .

هذا التحويل هو جمع لعمليتي الازاحة translation والدوران rotation للعنصر المدروس أي يتم ازاحة متوسط الاحداثيات x,y لكل البكسلات إلى مبدأ الاحداثيات و يتم تدوير الصورة بزاوية بحيث (وبشكل هندسي) يتم توجيه الصورة المحولة بالاتجاه الذي تبدوا فيه متطاولة أكثر .

يعتبر هذا التحويل مفيداً في مجال التعرف على الصور image recognition بما أن شخصية الشيء المدروس غير معروفة وربط الصورة مع محاورها الأساسية يمكن أن يساعد في ازالة تأثيرات الانزياح والدوران أثناء عملية التحليل.

ليكن لدينا k شعاع ذو n بعد :

$$X_1, X_2, \ldots\ldots, X_k$$

$$X=[\, X_1, X_2, \ldots\ldots, X_k]$$

مصفوفة التغاير covariance matrix للشعاع الاجمالي X الذي يضم كل الأشعة السابقة :

$$C_x = E[\,(X-m_x)(X-m_x)^T\,] \qquad (١٢-٢)$$

حيث m_x الشعاع المتوسط للأشعة X، C_x مصفوفة n*n حقيقية متناظرة فيها العنصر c_{ij} يساوي التغاير covariance بين الاحداثي i والاحداثي j للشعاع الاجمالي X. إذا كان $c_{ij}=0$ فإن الاحداثي i والاحداثي j غير مترابطين. وعندما يكون $c_{ij}>0$ أو فإن هناك ترابط بين الاحداثي i والاحداثي j .

ينتج عن تطبيق PCA على الشعاع الاجمالي X شعاع آخر Y (يتألف من k شعاع ذو n بعد) وفيه m_y و C_y الشعاع المتوسط ومصفوفة التغاير على التوالي للمصفوفة المحولة Y. لكن وكما سنرى لاحقاً يكون $m_y=0$ و C_y مصفوفة قطرية n*n مما يعني أن الاحداثي i والاحداثي j للشعاع الاجمالي Y غير مترابطين من أجل كل $i\neq j$.

لفحص أثر استخدام PCA في ربط صورة ثنائية الأبعاد مع محاورها الأساسية نعتبر أنه لدينا صورة m*n بكسل. إذا تم تمثيل الاحداثيات (x,y) لكل بكسل بشعاع ذو بعدين فإن الصورة يتم تمثيلها بمجموعة m*n من أشعة ثنائية الأبعاد. عند تطبيق PCA على هذا الشعاع الاجمالي فإن الشعاع المتوسط للشعاع الاجمالي الناتج عن التحويل $m_y=0$ ومصفوفة التغاير للشعاع الاجمالي الناتج ستكون مصفوفة قطرية مما يعني أن الاحداثيين x و y غير مترابطين.

وهذا يعني هندسياً أن متوسط الصورة الأصلية تمت ازاحته إلى مبدأ الاحداثيات وتم تدوير الصورة بحيث يتم ربطها مع محاورها الأساسية.

٢-٢-٢-٤ استخدام PCA لتصحيح الوضعية pose correction للوجه ثلاثي الأبعاد :

بعد اقتطاع الوجه والتخلص من الضجيج نصحح وضعيته باستخدام PCA ثم نصحح الأخطاء المتبقية بعد التصحيح. لتكن P مصفوفة n*٣ للإحداثيات x,y,z لمجموعة النقاط الممثلة للوجه:

$$P = \begin{pmatrix} x1\ x2\ \ldots\ xn \\ y1\ y2\ \ldots\ yn \\ z1\ z2\ \ldots\ zn \end{pmatrix}$$

الشعاع الوسطي ومصفوفة التغاير للمصفوفة P يعطيان بالعلاقتين التاليتين:

$$m = ١/n \sum^{n}_{k=١} P_k \qquad (١٣-٢)$$

$$C = ١/n \sum^{n}_{k=١} P_k . P_k^T - m.m^T \qquad (١٤-٢)$$

حيث P_k العمود k للمصفوفة P. مصفوفة الأشعة الخاصة V لمصفوفة التغاير C تعطى بالعلاقة:

$$CV=DV \qquad (١٥-٢)$$

حيث D مصفوفة القيم الخاصة لC. يمكن ربط P مع محاورها الأساسية باستخدام العلاقة التالية والتي تمثل PCA:

$$P_١=V^T.(P-m) \qquad (١٦-٢)$$

تكرر العملية أعلاه حتى تتقارب V إلى المصفوفة الواحدية حيث :

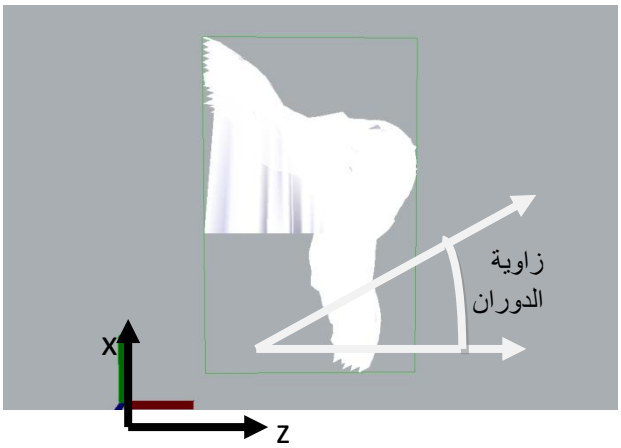
V^T تمثل مصفوفة الدوران rotation matrix وتستخدم لتصحيح الوضعية. m تمثل متوسط الاحداثيات وهي تمثل شعاع الإزاحة translation vector وتستخدم لإزاحة متوسط مجموعة النقاط إلى مبدأ الاحداثيات وبالتالي تتألف العملية من إزاحة translation ودوران rotation.

٢-٢-٢-٥ تصحيح الأخطاء الباقية بعد عملية التصحيح باستخدام PCA:

توضيح عملية التصحيح بالنسبة لزوايا الدوران المختلفة:

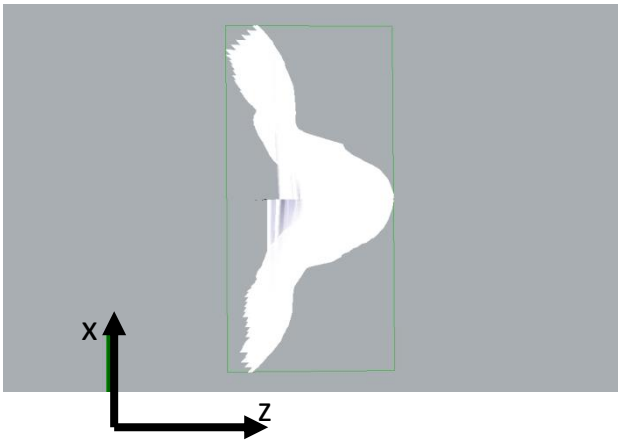
توضح الأشكال التالية مساقط للوجه ثلاثي الأبعاد على المستويات الثلاثة وحالات الدوران ومسقط الوجه على المستوي نفسه بعد التصحيح بالاستفادة من PCA:

الشكل التالي يبين مسقط للوجه ثلاثي الأبعاد على المستوي x,z حيث يوضح وجود دوران حول المحور y



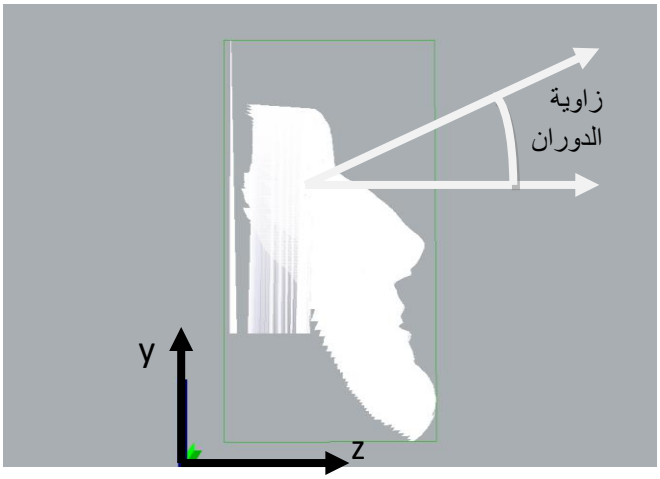
الشكل (٢-٥) المسقط على المستوي x,z قبل عملية التصحيح

الشكل التالي يبين مسقط الوجه السابق على المستوي x,z بعد عملية التصحيح باستخدام PCA حيث نلاحظ أنه بعد تصحيح الدوران يصبح التبعثر أعظمياً وفق المحور x وتقليل الترابط بين المتغيرين x,z.



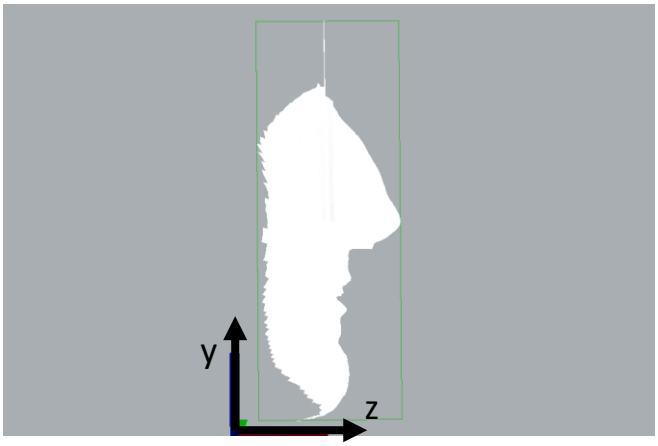
الشكل (٢-٦) المسقط على المستوي x,z بعد عملية التصحيح

الشكل التالي يبين مسقط للوجه ثلاثي الأبعاد على المستوي y,z حيث يوضح وجود دوران حول المحور x :



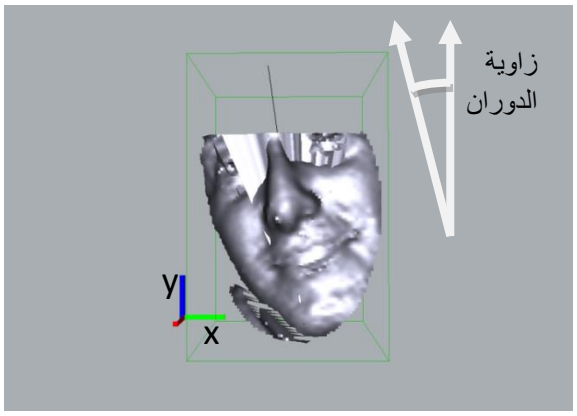
الشكل (٢-٧) المسقط على المستوي y,z قبل عملية التصحيح

الشكل التالي يبين مسقط الوجه السابق على المستوي y,z بعد عملية التصحيح باستخدام PCA حيث نلاحظ أنه بعد تصحيح الدوران يصبح التبعثر أعظمياً وفق المحور y وتقليل الترابط بين المتغيرين y,z :



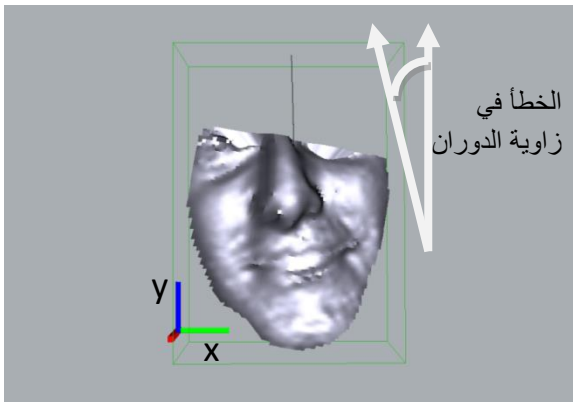
الشكل (٢-٨) المسقط على المستوي y,z بعد عملية التصحيح

الشكل التالي يبين الوجه ثلاثي الأبعاد وفق المستوي x,y حيث يوضح وجود دوران حول المحور z:



الشكل (٢-٩) يوضح وجود زاوية دوران حول المحور z قبل عملية التصحيح

الشكل التالي يبين الوجه السابق وفق المستوي x,y بعد محاولة عملية التصحيح باستخدام PCA حيث نلاحظ أنه بعد محاولة تصحيح الدوران يصبح التبعثر أعظمياً وفق المحور y وتقليل الترابط بين المتغيرين x,y:



الشكل (٢-١٠) يوضح وجود خطأ زاوية دوران حول المحور z بعد عملية التصحيح باستخدام PCA

نلاحظ من الشكل السابق أن هذه الطريقة قد تعطي أخطاء دوران حول المحور z بعد التصحيح باستخدام PCA والسبب في ذلك هو أن شكل الوجه بعد الاقتطاع يجب أن يكون بيضوياً مثالياً محوره الأكبر وفق مستقيم حافة الأنف في المستوي x,y ومركزه قمة الأنف كي تعطي عملية التصحيح أفضل

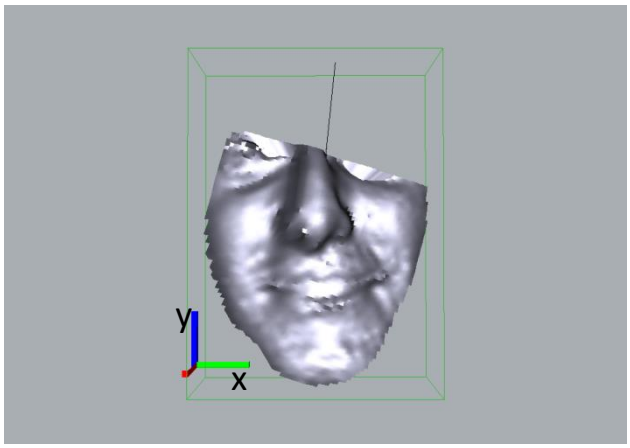
نتيجة لأن PCA تؤدي إلى جعل الشكل بيضوي في المستوي x,y متطاول وفق المحور y (التبعثر أعظمي وفق المحور y)

لكن وبسبب التغيرات في وضعية الوجه وظروف التقاط الصورة نحصل بعد عملية الاقتطاع على شكل بيضوي غير مثالي(محوره الأكبر يميل عن مستقيم حافة الأنف) وبالتالي نتيجة استخدام PCA للتصحيح هي استمرار الحصول على دوران حول المحور z

لمعالجة هذه المشكلة وبعد عملية التصحيح باستخدام PCA نحدد النقاط التي تشكل مستقيم حافة الأنف عن طريق تحديد النقاط المرشحة لتكون قمة الأنف في شرائح الوجه الأفقية التي تكون في مستوى أعلى وفق المحور y من مستوى قمة الأنف المحددة سابقاً وهذه النقاط التي حددناها تشكل حافة الأنف الذي يشكل مستقيماً يجب أن يكون ذو ميل لانهائي بعد التصحيح في الحالة المثالية(لا يوجد خطأ دوران للوجه حول المحور z)

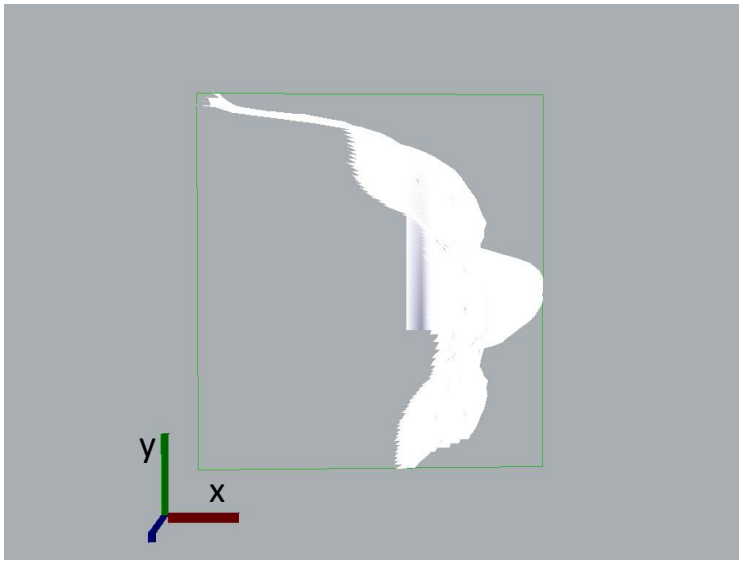
وبالتالي لمعالجة وجود خطأ دوران حول المحور z بعد التصحيح باستخدام PCA نقوم بإجراء دوارن حول المحور z للوجه عندما تنقص قيمة ميل مستقيم حافة الأنف عن ١٤,٣٠١ درجة أي عندما يتجاوز الخطأ في زاوية الدوران ٤ درجات وسبب أخذ هامش ٤ درجات هو بسبب عدم امكانية التحديد الدقيق لميل هذا المستقيم وبالتالي نأخذ هامش ٤ درجات قبل أن نعتبر أن هناك دوران واضح لهذا المستقيم في المستوي x,y (وبالتالي دوران للوجه حول المحور z).

الشكل التالي يبين الوجه السابق وفق المستوي x,y بعد عملية التصحيح باستخدام PCA وبعد اجراء عملية تصحيح خطأ الدوران حول المحور z وفق الأسلوب السابق:

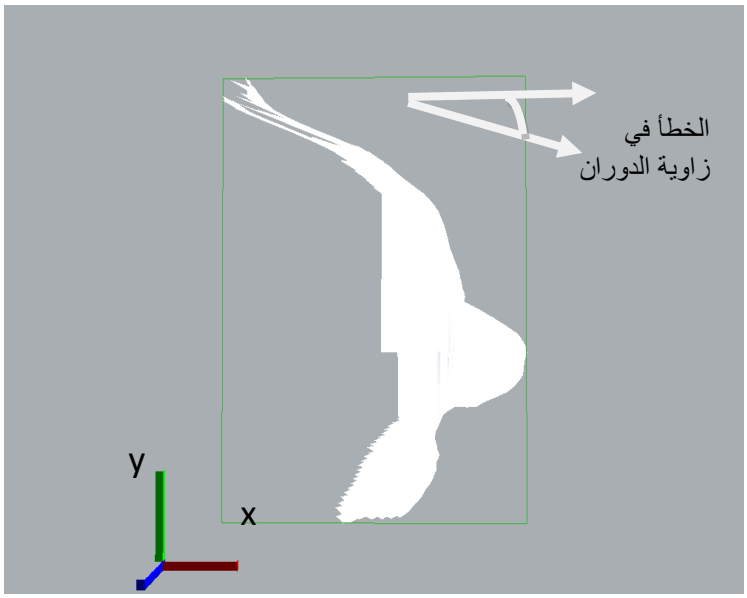


الشكل (٢-١١) الوجه ثلاثي الأبعاد وفق المستوي x,y بعد عملية التصحيح باستخدام PCA وبعد تصحيح خطأ زاوية الدوران حول المحور z

نلاحظ أن هذه الأسلوب في تصحيح الوضعية يواجه مشكلة أخرى عندما يكون هناك اقتطاع غير متناظر للوجه بالنسبة للأنف (مثلاً بسبب تغيرات الوضعية – قد تكون زاوية الدوران حول المحور y للوجه ثلاثي الأبعاد كبيرة نسبياً) مما يؤدي إلى الحصول على وجه بعد التصحيح لا يزال فيه دوران حول المحور y بزاوية معينة:

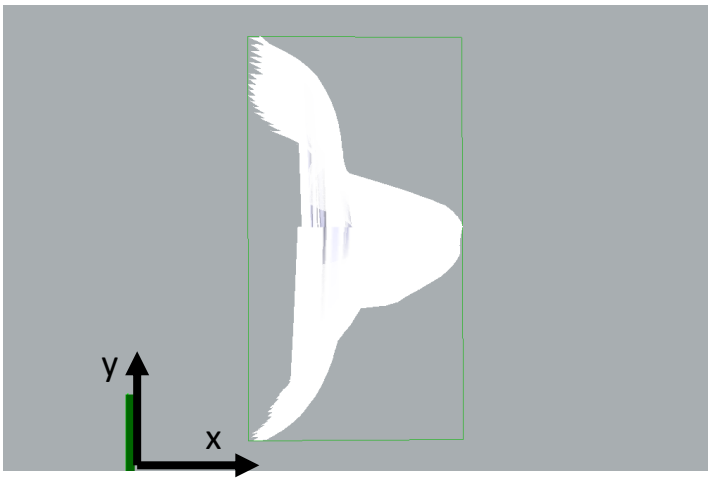


الشكل (٢-١٢) اقتطاع غير متناظر لمنطقة الوجه بالنسبة للأنف



الشكل (٢-١٣) نتيجة التصحيح بتطبيق مرحلة واحدة من الاقتطاع والتصحيح

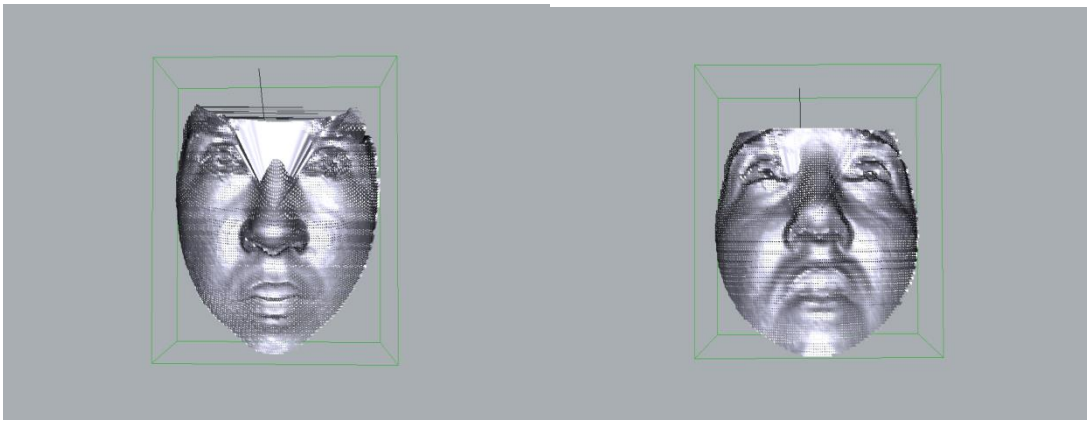
وعالجنا ذلك بتطبيق مرحلة أخرى من الاقتطاع ثم تصحيح الوضعية(مرحلة معالجة أولية ثانية) حيث مرحلة الاقتطاع الثانية ستقتطع من الوجه بشكل متناظر تقريباً حيث في مرحلة الاقتطاع الأولى يتم اقتطاع منطقة بيضوية أبعادها ٦٥mm وفق المحور x و ٨٥mm وفق المحور y ثم في مرحلة الاقتطاع الثانية يتم اقتطاع منطقة بيضوية أبعادها ٦٠mm وفق المحور x و ٧٥mm وفق المحور y:



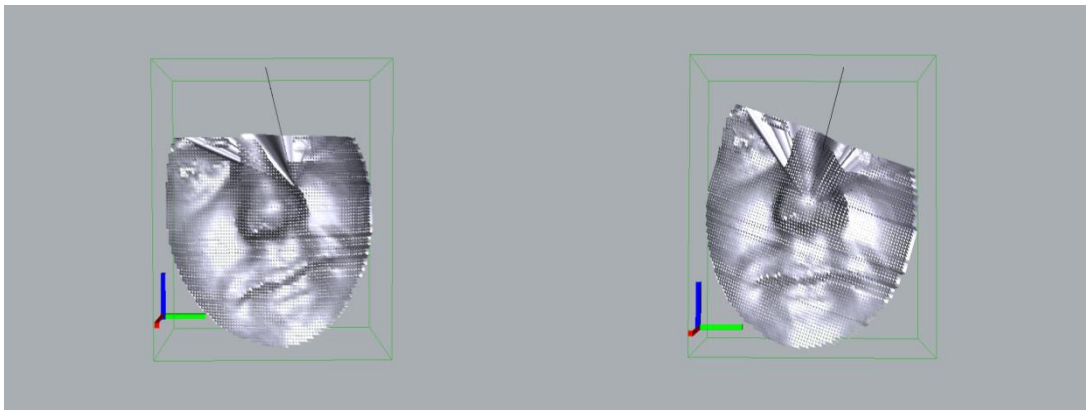
الشكل (٢-١٤) نتيجة التصحيح بتطبيق مرحلتين من الاقتطاع والتصحيح

٢-٢-٢-٦ نتائج مرحلة تصحيح الوضعية:

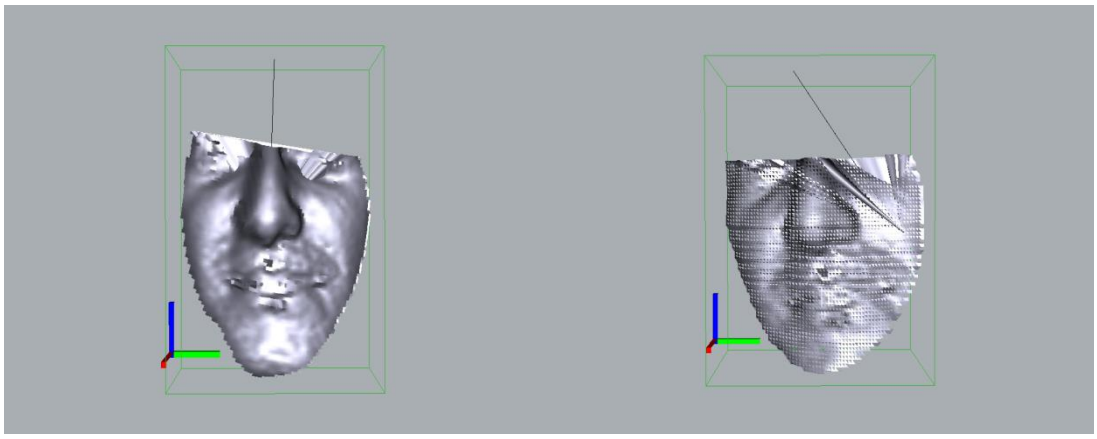
نوضح في الأشكال التالية أمثلة عن وجوه ثلاثية الأبعاد قبل وبعد تصحيح الوضعية :



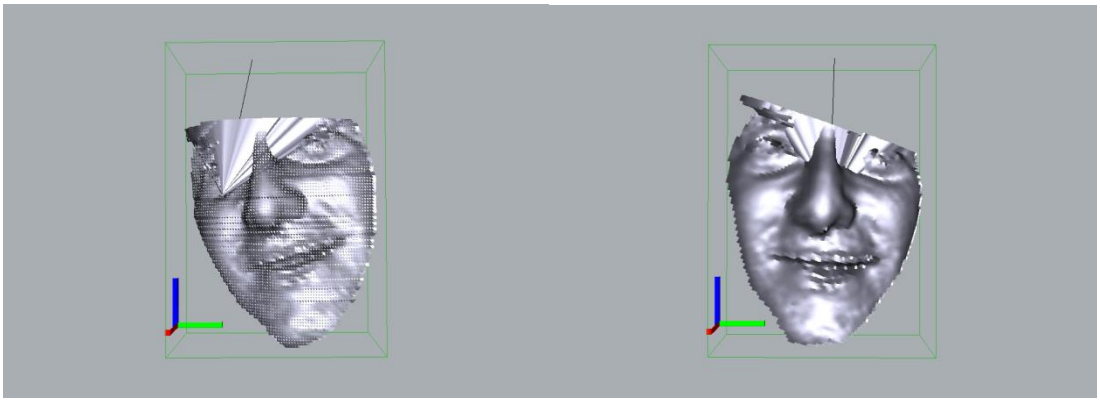
الشكل(٢-١٥)الوجه قبل التصحيح(دوران حول المحور x) الوجه بعد التصحيح



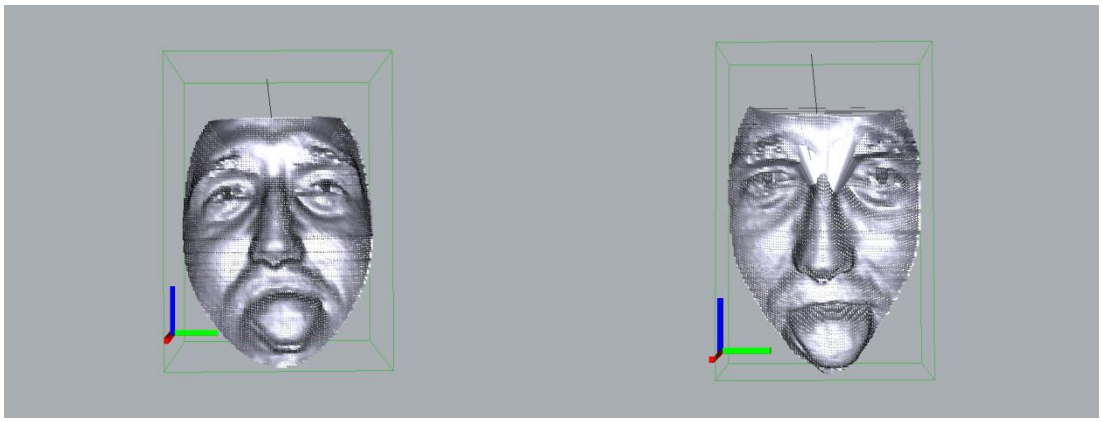
الشكل(٢-١٦)الوجه قبل التصحيح(دوران حول المحور z) الوجه بعد التصحيح



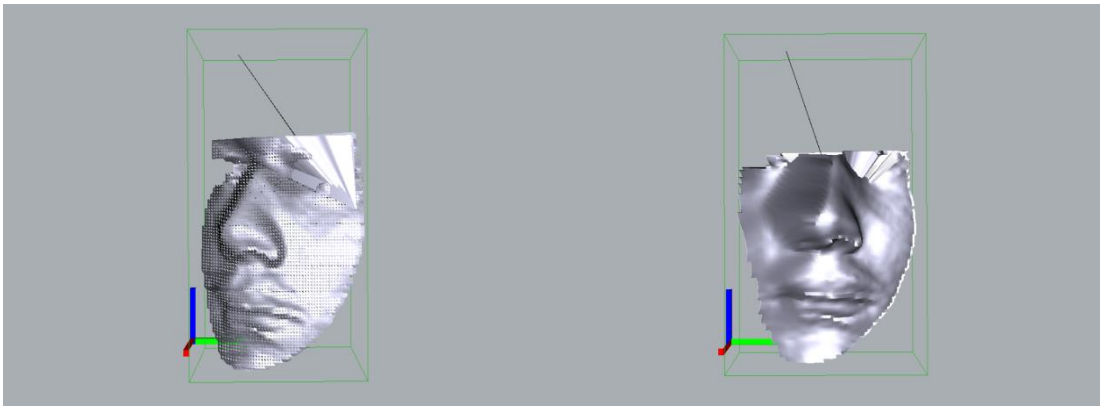
الشكل(٢-١٧)الوجه قبل التصحيح(دوران حول المحور y) الوجه بعد التصحيح



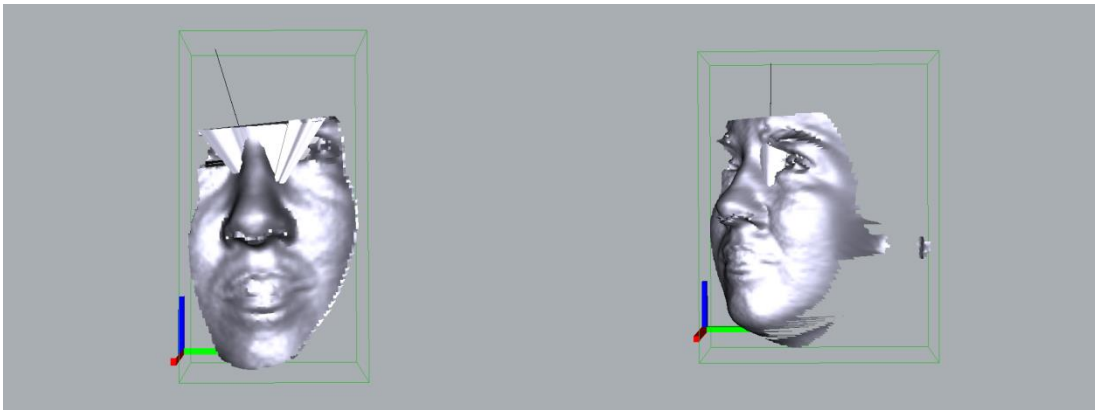
الشكل(٢-١٨)الوجه قبل التصحيح(دوران حول المحورين y,z) الوجه بعد التصحيح



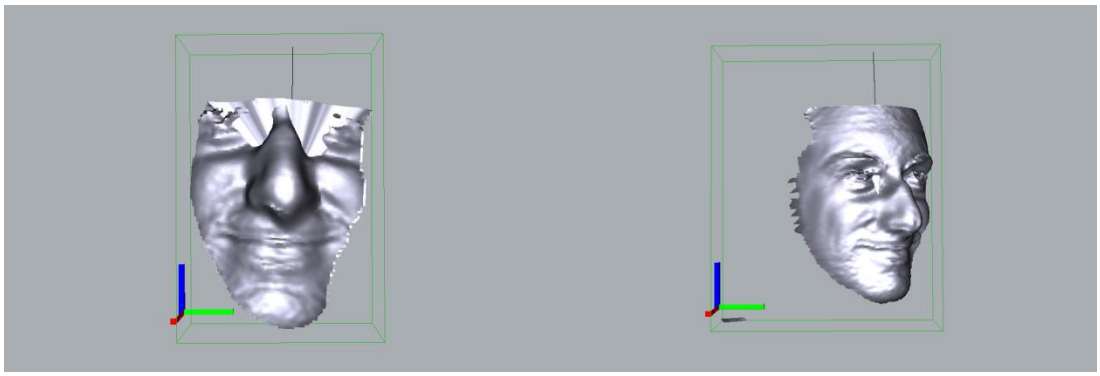
الشكل(٢-١٩)الوجه قبل التصحيح(دوران حول المحور x) الوجه بعد التصحيح



الشكل(٢-٢٠)الوجه قبل التصحيح(دوران حول المحور y) الوجه بعد التصحيح



الشكل(٢-٢١)الوجه قبل التصحيح (دوران حول المحور y) الوجه بعد التصحيح

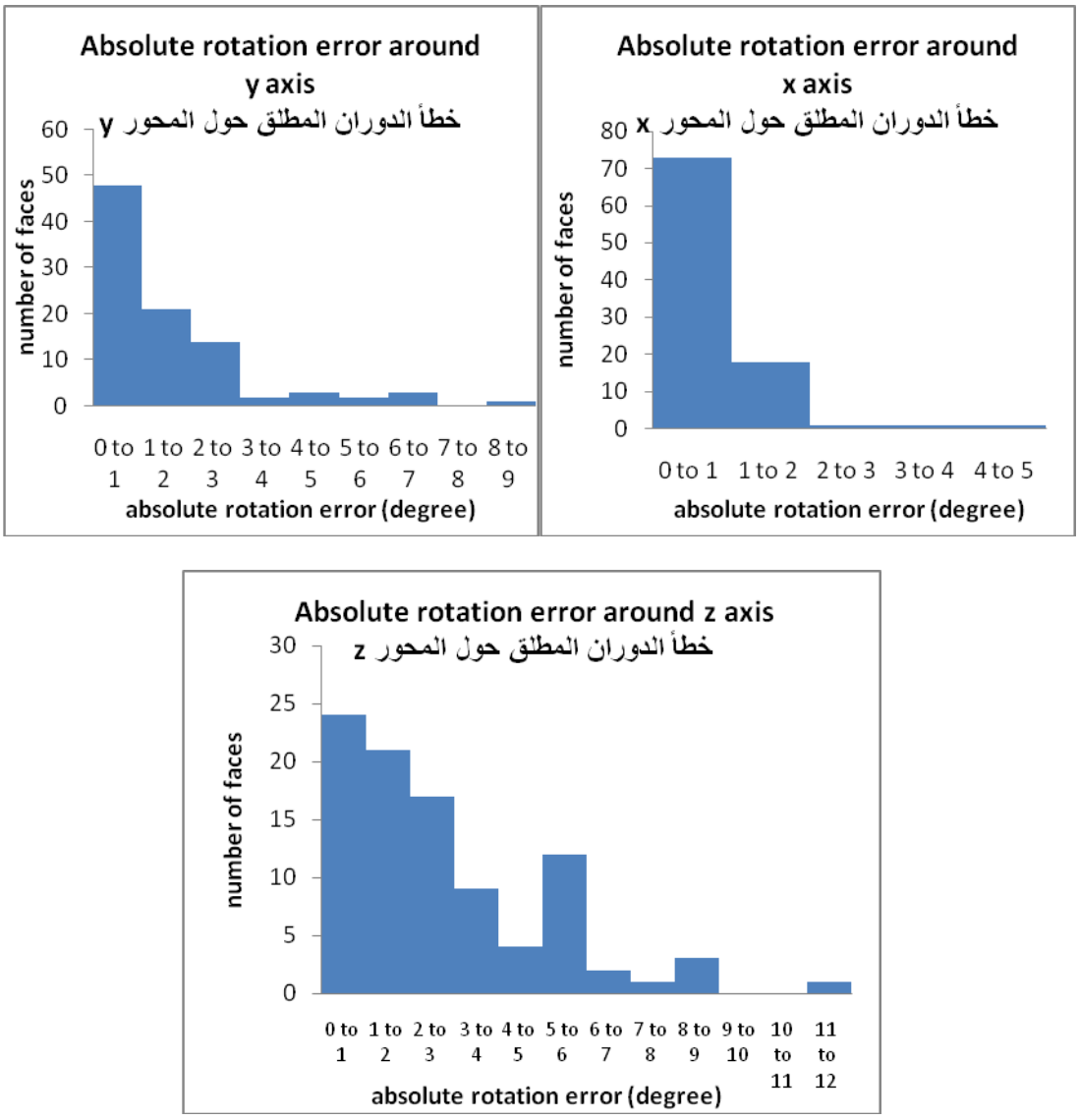


الشكل(٢-٢٢)الوجه قبل التصحيح(دوران حول المحور y) الوجه بعد التصحيح

تغيرات الوضعية pose تقع في حدود زوايا الدوران التالية :

(٤٥ درجة حول المحور y ، ٣٥ درجة حول المحور x) وهي زوايا الدوران التي من أجلها تعمل خوارزمية تصحيح الوضعية pose correction وهي الزوايا التي يتحقق من أجلها التحديد الصحيح لقمة الأنف. سبب العمل ضمن حدود تغيرات الزوايا المذكورة أعلاه هو أن زيادة الزاوية إلى ما يقترب من ٤٥ درجة وأكثر سينتج عنه بعد عملية التصحيح زيادة في زاوية الدوران إلى ٩٠ درجة. وبما أننا استخدمنا طريقة تصحيح الدوران حول المحور z بتصحيح ميل مستقيم حافة الأنف في المستوي x,y فإن ذلك يعني امكانية تصحيح زاوية الدوران حول المحور z مهما كانت(في الحدود العملية) حيث تنخفض امكانية التصحيح عند زيادة زاوية الدوران بحيث تقترب من ٩٠درجة بسبب عدم امكانية التحديد الدقيق أو عدم امكانية التحديد مطلقاً لميل مستقيم حافة الأنف. ونلاحظ أن هذه الخوارزمية قوية تجاه التعابير في الوجه وفي حالة تغطية بعض أجزاء الوجه بواسطة اللسان.

ومن أجل القيام بالتحليل الكمي للنتائج يجب مقارنة وضعية كل وجه بعد التصحيح مع مرجع معين تكون فيه الوضعية أمامية دون دوران حول أحد المحاور لذلك نقارن الوجوه ثلاثية الأبعاد تحت الاختبار بعد التصحيح مع الوجه بالوضعية الأمامية والتي تعود للشخص نفسه بعد التصحيح والمستخدم في بناء قاعدة البيانات gallery وتتم المقارنة باستخدام خوارزمية ICP لحساب خطأي الانزياح والدوران ونوضح خطأ الدوران وخطأ الانزياح بينها في الأشكال التالية التي تبين في المحور الشاقولي عدد الحالات التي من أجلها حصلنا على خطأ بعد التصحيح يقع ضمن المجال المحدد في المحور الأفقي:

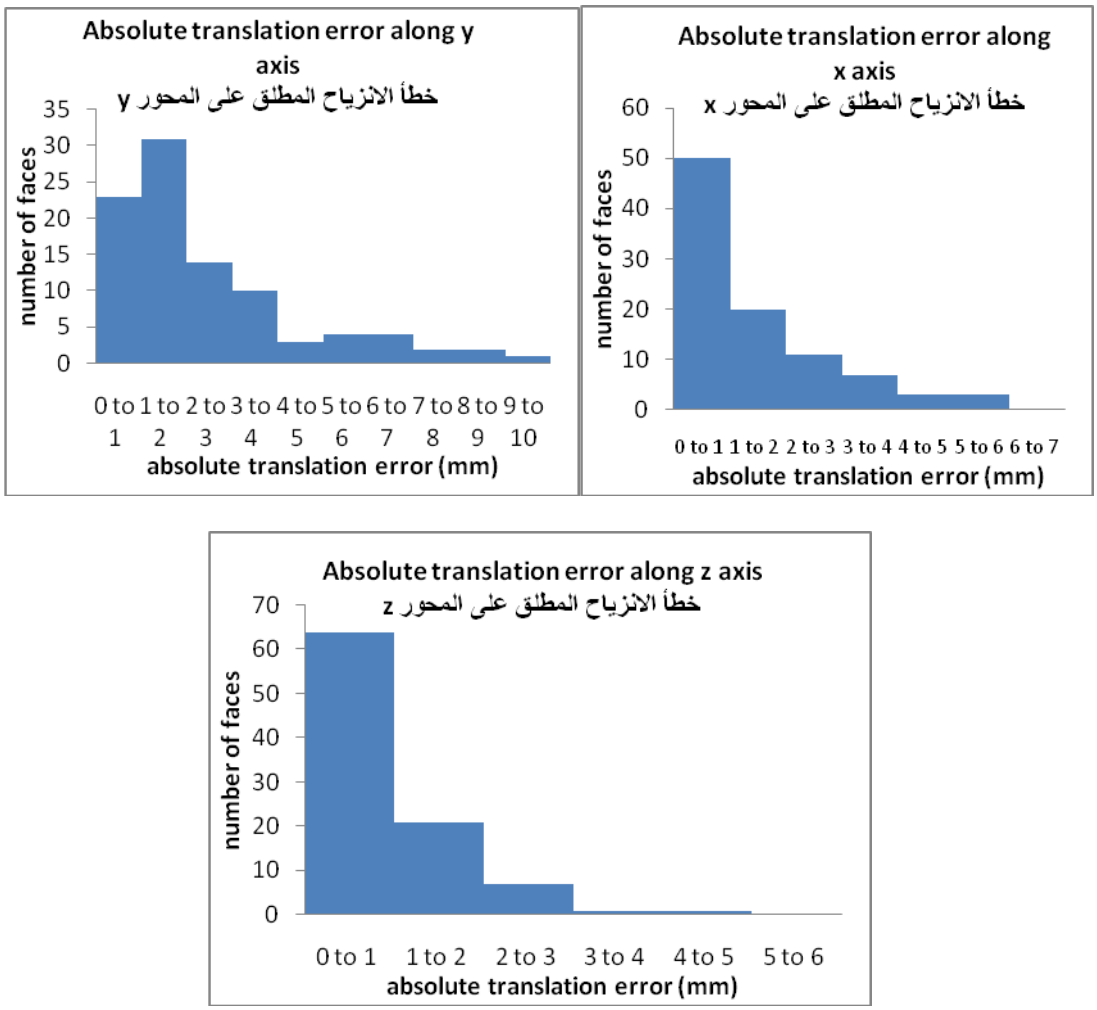


الشكل(٢-٢٣) خطأ الدوران

القيمة المتوسطة للخطأ المطلق في زاوية الدوران حول المحور x (٠,٦٥) درجة

القيمة المتوسطة للخطأ المطلق في زاوية الدوران حول المحور y (١,٥٢) درجة

القيمة المتوسطة للخطأ المطلق في زاوية الدوران حول المحور z (٢,٧١) درجة



الشكل (٢-٢٤) خطأ الانزياح

القيمة المتوسطة لخطأ الانزياح المطلق على المحور x (١,٤) mm

القيمة المتوسطة لخطأ الانزياح المطلق على المحور y (٢,٤٢) mm

القيمة المتوسطة لخطأ الانزياح المطلق على المحور z (٠,٨٤) mm

نلاحظ أن أكبر خطأ انزياح يكون على المحور y والسبب في ذلك هو احتمال وقوع خطأ في تحديد قمة الأنف وأخذ أحد النقاط التي تشكل حافة الأنف في المستويات الأفقية الأعلى قليلاً من مستوى قمة الأنف لتكون القمة المقترحة ويحدث ذلك أحياناً بسبب الاختلاف البسيط في أشكال أنوف بعض الأشخاص في ظروف معينة مثل تعابير معينة في الوجه وقد يؤدي ذلك إلى الحصول على مطال لقمة الأنف الفعلية أصغر قليلاً من مطال نقطة مجاورة تنتمي لحافة الأنف.

القيمة العظمى المتوسطة للخطأ المطلق في زاوية الدوران ٢,٧١ درجة والقيمة العظمى المتوسطة لخطأ الانزياح المطلق ٢.٤٢ mm

الفصل الثالث

التعرف الآلي على الوجوه ثلاثية الأبعاد

٣-١ التقنية المستخدمة للمقارنة والتعرف:

سوف نستفيد من خوارزمية النقطة الأقرب التكرارية(ICP) Iterative closest point algorithm.

في عملية المقارنة بين الوجوه ثلاثية الأبعاد. لاحظنا من نتائج عملية تصحيح الوضعية وجود خطأ انزياح ودوران بسيط عن الوضعية الأمامية النظامية وبالتالي للمقارنة بين هذه الوجوه ثلاثية الأبعاد لا بد من أسلوب يعوض هذه الأخطاء.

تم تقديم ICP من قبل Besl و McKay كأسلوب ناجح لمطابقة الأشكال بوجود انزياح بينها حيث يتألف الدخل من مجموعتي نقاط في فضاء ذو d بعد حيث ١=<d. المقياس النظامي للتشابه الذي تحاول الخوارزمية الوصول إلى قيمته المثالية الدنيا هو:

(mean squared distance) أي متوسط مربع المسافة بين النقاط الأقرب في الشكلين.

تعتبر عملية تحليل ومطابقة الأشكال والنماذج عملية مهمة لكثير من التطبيقات وخصوصاً في مجال الابصار الحاسوبي computer vision والتعرف على النماذج pattern recognition. المرجع[٩]

عادة يكون لدينا نموذجين A و B ونريد أن نحدد كم يشبه هذين النموجين بعضهما وعادة أحد هذين النموذجين قد يخضع لعمليات تحويل transformation محددة مثل الانزياح translation والدوران rotation بهدف مطابقته مع النموذج الآخر بشكل جيد.

في عدة حالات يتم تمثيل النماذج بمجموعات منتهية من النقاط في فضاء ثنائي أو ثلاثي الأبعاد (وعندئذ يشار إليها "نماذج نقاط"). لقياس "التشابه" بينهما يمكن استخدام"توابع كلفة cost function" متنوعة وأحد هذه التوابع المستخدمة بشكل شائع هو مجموع مربع المسافات أو متوسط مربع المسافات وتحت هذا الاعتبار يكون تابع الكلفة :

$$\Phi_2(A,B) = \frac{1}{m}\sum_{a\in A} \|a - N_B(a)\|^2 \quad (٣-١)$$

حيث $N_{B(a)}$ تمثل الجار الأقرب nearest neighbour للنقطةa في النموذج B وحيث m=|A|

وفي العلاقة التالية نستخدم التعبير التالي :

$$RMS(t) := \frac{1}{m} \sum_{a \in A} \|a + t - N_B(a + t)\|^2 \qquad (٢-٣)$$

تعتبر خوارزمية النقطة الأقرب التكرارية (خوارزمية ICP) خوارزمية مطابقة شائعة الاستخدام بسبب بساطتها وأدائها الجيد.

إذا كان لدينا مجموعتي نقاط A وB في الفضاء R_d ونرغب بتصغير تابع كلفة $\Phi(A+t,B)$ بواسطة الانزياحات t لA بالنسبة لB. تبدأ الخوارزمية بانزياح يطابق A مع B جزئياً وبعدئذ تقوم بشكل متكرر بخطوات محلية متتالية تحسن تطابق A مع B من حيث المكان والوضعية في الفضاء R_d مع انقاص تابع الكلفة المعطى $\Phi(A+t,B)$ حتى نصل إلى حالة الاستقرار وهذا يتم كما يلي:

في العملية رقم i من هذه العمليات المتكررة في خوارزمية ICP تكون مجموعة النقاط A قد تم ازاحتها بشعاع $t_{i-١}$ حيث $\vec{t_{٠}} = ٠$.

بعد ذلك نطبق الخطوتين التاليتين :

١) نسند إلى كل نقطة مزاحة $a+t_{i-١} \in A+t_{i-١}$ إلى أقرب جار $b=N_B(a+t_{i-١}) \in B$ حسب المسافة الديكارتية

٢) بعد ذلك نحسب الانزياح النسبي الجديد Δti الذي يقلل تابع الكلفة Φ (بالنسبة لمجموعة النقاط المتقابلة في هذه المرحلة) وبشكل أكثر تحديداً نوجد Δti التي تقلل :

$$\phi_2(A + t_{i-1}, \Delta_{t_i}, B) = \frac{1}{m} \sum_{a \in A} \|a + t_{i-1} + \Delta t_i - N_B(a + t_{i-1})\|^2 \qquad (٣-٣)$$

بعد ذلك نربط نقاط A مع نقاط B بتطبيق انزياح على هذه النقاط بمقدار Δti وبالتالي يصبح الانزياح الكامل $t_i = t_{i-١} + \Delta t_i$.

تقوم خوارزمية ICP بتطبيق هاتين الخطوتين بشكل متكرر وتتوقف عندما تثبت قيمة تابع الكلفة وتتوقف عن التناقص بالنسبة للمرحلة السابقة وعملياً نعتبر أننا وصلنا إلى هذه الحالة عندما يصبح الفرق في قيمة تابع الكلفة أقل من عتبة معينة $\tau > 0$)

بكلمات أخرى: في الخطوة ١ لكل عملية تكرار في خوارزمية ICP نسند النقاط في النسخة المزاحة الحالية لA إلى مجاوراتها الأقرب في B وفي الخطوة ٢ نقوم بإزاحة نقاط A بهدف تقليل قيمة تابع الكلفة بالنسبة لمجموعة النقاط المتقابلة الحالية والمحسوبة في الخطوة رقم ١ وهذا قد يتسبب بدوره في اسناد نقاط مجاورة أقرب جديدة في B لبعض النقاط في النسخة المزاحة الجديدة من A مما يجعل الخوارزمية تقوم بالمزيد من هذه العمليات المتكررة وفق الخطوتين السابقتين من أجل مجموعة النقاط المتقابلة الجديدة. إذا لم تغير أي نقطة من A جارتها الأقرب في B فإن قيمة تابع الكلفة لن تتغير في عملية التكرار التالية وبالتالي تتوقف الخوارزمية.

نلاحظ أن عملية مطابقة النماذج Pattern matching التي تقوم بها هذه الخوارزمية وحيدة الاتجاه one-directional أي أنها تسعى لإيجاد انزياح لA يضع نقاط A قرب نقاط B وليس بالعكس.

أحد المشاكل الأساسية في هذا الأسلوب للتعرف على الوجوه ثلاثية الأبعاد هي أنها مكلفة حسابياً أي أن عملية مقارنة واحدة بين زوجين من سطحي الوجهين ثلاثيي الأبعاد قد تأخذ وقتاً طويلاً يتناسب مع عدد النقاط في الوجهين ثلاثيي الأبعاد ومع سرعة المعالج في الحاسب المستخدم. وبالتالي تصبح العملية غير مجدية عند البحث في قواعد بيانات كبيرة. وهذا الجهد الحسابي الكبير يأتي من عاملين:

الأول: إن حساب MSE(mean squared error) أي متوسط مربع الخطأ بين أزواج النقاط الأقرب في السطحين يحتاج عمليات حسابية من مرتبة O(MN) حيث M عدد النقاط ثلاثية الأبعاد على أحد السطحين وN عدد النقاط على السطح الآخر.

الثاني: تحتاج أيضاً كل عملية مقارنة عمليات مطابقة للموضع والوضعية متكررة بطيئة لربط زوجين من سطوح الوجوه ويمكن لهذه العملية التكرارية أن تنتهي بشكل خاطئ إلى نهاية محلية.

وهناك مشكلة أخرى في هذا الأسلوب المقترح للتعرف على الوجوه ثلاثية الأبعاد وهي أنها ليست قوية تجاه التغيرات في تعابير الوجه.

٣-٢ تسريع عملية التعرف باستخدام تكمية للتمثيل ثلاثي الأبعاد للوجه:

تسريع عملية التعرف هي العملية التي تستبعد منذ البداية وقبل القيام بعملية التعرف البطيئة نسبة كبيرة من الأصناف المرشحة وإذا كانت M تمثل حجم قاعدة البيانات فإن:

كفاءة الاستبعاد Eff هو العدد المتوقع للأصناف المستبعدة مقسومة على العدد الكلي للأصناف M في قاعدة البيانات.

وكلما كبرت قيمة Eff تعتبر عملية الاستبعاد أفضل (تسريع أكبر لعملية التعرف). لم يكن هناك مفر من استخدام هذا الأسلوب لتسريع عملية التعرف في هذا البحث لأن حجم البيانات كبير فهناك ٣٢ وجه في قاعدة البيانات وهذا يعني أن هناك ٣٢ عملية مقارنة عند التعرف على وجه تحت الاختبار وخاصة أن خوارزمية ICP المستخدمة في هذا البحث لعلمية المقارنة والتعرف تأخذ وقتاً نوعاً ما بينما باستخدام عملية استبعاد ذات كفاءة Eff=٠.٧٥ مثلاً فإن عدد عمليات المقارنة سينقص إلى ٨=٠,٢٥*٣٢ .

حتى تتحقق عملية التسريع يجب البحث عن تمثيل آخر للوجوه ثلاثية الأبعاد يخفض من حجم البيانات (يقلل عدد النقاط) الممثل لكل وجه لأن عملية البحث عن أزواج النقاط الأقرب في مجموعتي نقاط كبيرتين أثناء عملية المقارنة والتعرف تستهلك وقتاً يتناسب مع حجم البيانات الممثلة للوجوه(مع عدد النقاط في الوجوه ثلاثية الأبعاد).

وفي نفس الوقت يجب أن يحافظ هذا التمثيل الجديد للوجوه على نسبة تمثيل كافية للوجوه يحافظ على وصف كافي للوجه كي يبقى متميزاً عن الوجوه الأخرى.

يبين الشكل (٣-٢) المخطط الصندوقي لخوارزمية التعرف على الوجوه ثلاثية الأبعاد بشكل كامل والتي تضم عملية الاستبعاد .

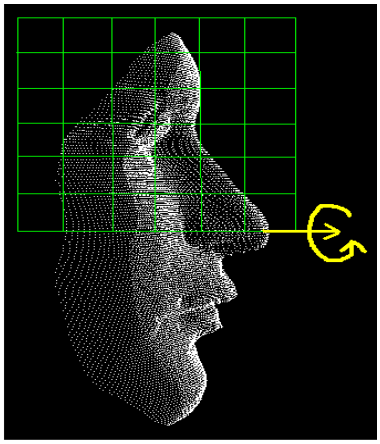
إذا اقترحنا عملية تكمية quantization للوجه ثلاثي الأبعاد وفق دوائر متمركزة عند قمة الأنف وقارناها مع تمثيل آخر وهو التمثيل الدائري spin image للوجه الذي نحصل عليه بالدوران حول محور يمر بنقطة قمة الأنف وجمع نقاط الوجه التي تمر عبر الدوائر في الوجه(الشكل ٣-١).

فإننا نجد في البداية أن التمثيل الدوراني spin image للوجه يصف الوجه بشكل أفضل من التمثيل بواسطة دوائر متمركزة حول قمة الأنف. ونعلم أن أداء عملية التعرف وفق طريقة تمثيل معينة يرتبط

مباشرة بمقدار وصفها للوجه ولكن نعلم أيضاً أنه كلما زادت كمية وصف الوجه وفق طريقة تمثيل معينة فإنه يصبح أكثر حساسية لتعابير الوجه.

وأيضاً إذا قارنا حسب التعقيد الرياضي فإن التمثيل الدوراني للوجه أكثر تعقيداً من التمثيل وفق الدوائر المتمركزة لذلك يعتبر من الأفضل استخدام هذا الأسلوب الأخير.

نحسب تمثيل الدوائر المتمركزة للوجوه في قاعدة البيانات و نخزنها لتشكل قاعدة بيانات أخرى وفق هذا التمثيل للوجوه ثم تتم عملية المقارنة باستخدام خوارزمية النقطة الأقرب التكرارية ICP بين هذا التمثيل للوجوه في قاعدة البيانات مع تمثيل الدوائر المتمركزة المقابل للوجه تحت الاختبار المراد معرفة شخصيته. وينتج عن عملية المقارنة شعاع قيم تشابه ذو حجم M (حيث M حجم قاعدة البيانات) وحسب نسبة الكفاءة أو نسبة الاستبعاد نقوم باختيار الوجوه في قاعدة البيانات ذات الخطأ الأقل بحيث نحصل على الوجوه المرشحة لتحقق التطابق مع الوجه تحت الاختبار.

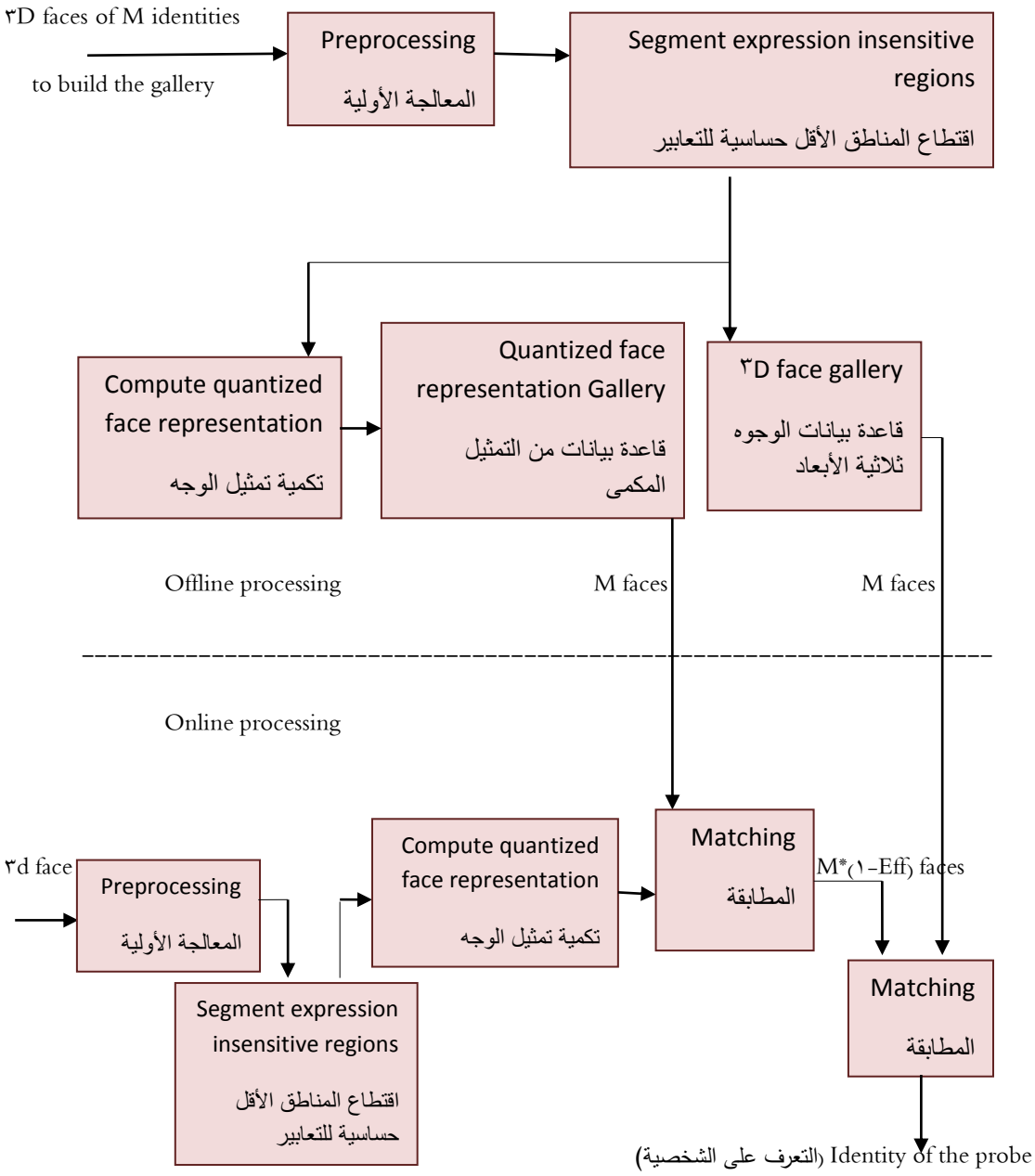


الشكل (٣-١) التمثيل الدوراني spin image

هناك مقايضة بين نسبة الكفاءة لعملية الاستبعاد (سرعة عملية التعرف) ودقة عملية التعرف بسبب التكمية لتمثيل الوجه(تمثيل غير كامل) فكلما زادت نسبة الكفاءة (أي كلما زادت سرعة التعرف) نقصت دقة التعرف والعكس بالعكس.

٣-٣ تجزيء الوجه والتعرف عليه: face segmentation and recognition

يبين الشكل (٣-٢) المخطط الصندوقي للخوارزمية الكاملة التي نقدمها للتعرف على الوجوه ثلاثية الأبعاد بما فيها عملية الاستبعاد وعملية التعرف النهائية من أجل حالة التجزيء المنتظم للوجوه:



الشكل (٣-٢) المخطط الصندوقي للخوارزمية الكاملة التي نقدمها للتعرف على الوجوه ثلاثية الأبعاد

تشكيل قاعدة البيانات:

خلال مرحلة المعالجة offline نشكل قاعدة البيانات من صور الوجوه ثلاثية الأبعاد. وبالأصل يتم استخدام وجه واحد ثلاثي الأبعاد(لا يحوي تعابير في العادة) لكل شخص. يتم معالجة كل وجه ثلاثي الأبعاد كما هو موضح أعلاه.

على الرغم من أن التعرف على الوجوه ثلاثية الأبعاد تحقق معدلات تعرف أعلى من التعرف على الوجوه ثنائية الأبعاد فإنها حساسة أكثر تجاه تعابير الوجه حيث ينخفض معدل التعرف من أجل الوجوه ثلاثية الأبعاد التي تحوي تعابير عند مقارنتها مع الوجوه ثلاثية الأبعاد والتي لا تحوي تعابير والمستخدمة في بناء قاعدة البيانات.

للتغلب على هذه المشكلة نجزء الوجوه ثلاثية الأبعاد إلى مناطق حساسة وغير حساسة(أقل تأثر) تجاه التعابير. هناك طريقتان تستخدمان لهذا الغرض وكلا الطريقتين أتوماتيكية تماماً لكن الطريقة الأولى تجزء الوجوه بشكل غير منتظم بالاعتماد على خواص كل وجه على حدى بينما الطريقة الثانية تقوم بالتجزيء المنتظم أي يتم اقتطاع نفس الملامح من كل الوجوه.

هذه المناطق المقتطعة ستشكل قاعدة البيانات من الوجوه المعروفة للنظام مما يقلل من تأثير التغيرات الناتجة عن التعابير في الوجوه تحت الاختبار.

تشكيل قاعدة البيانات من تمثيل الدوائر المتمركزة:

لو حسبنا تمثيل الدوائر المتمركزة من الوجه المستخدم في بناء قاعدة البيانات قبل اقتطاع المناطق الأقل تأثراً بالتعابير فإنه ستزيد الأخطاء بالنسبة لوجوه الاختبار التي تحوي تعابير عند مقارنة تمثيل الدوائر المتمركزة لها مع تمثيل الدوائر المتمركزة للوجوه في قاعدة البيانات والنتيجة معدل تعرف غير مقبول. حيث يصل معدل التعرف على الوجوه بوجود تعابير إلى (RR=٨١,٢٥%)من أجل تجزيء منتظم للوجوه من أجل نسبة استبعاد ٧١,٨٧٥%.

لتخفيف تأثير التعابير في الوجوه تحت الاختبار والمحافظة على دقة ومعدل تعرف مقبول لدينا حلان:

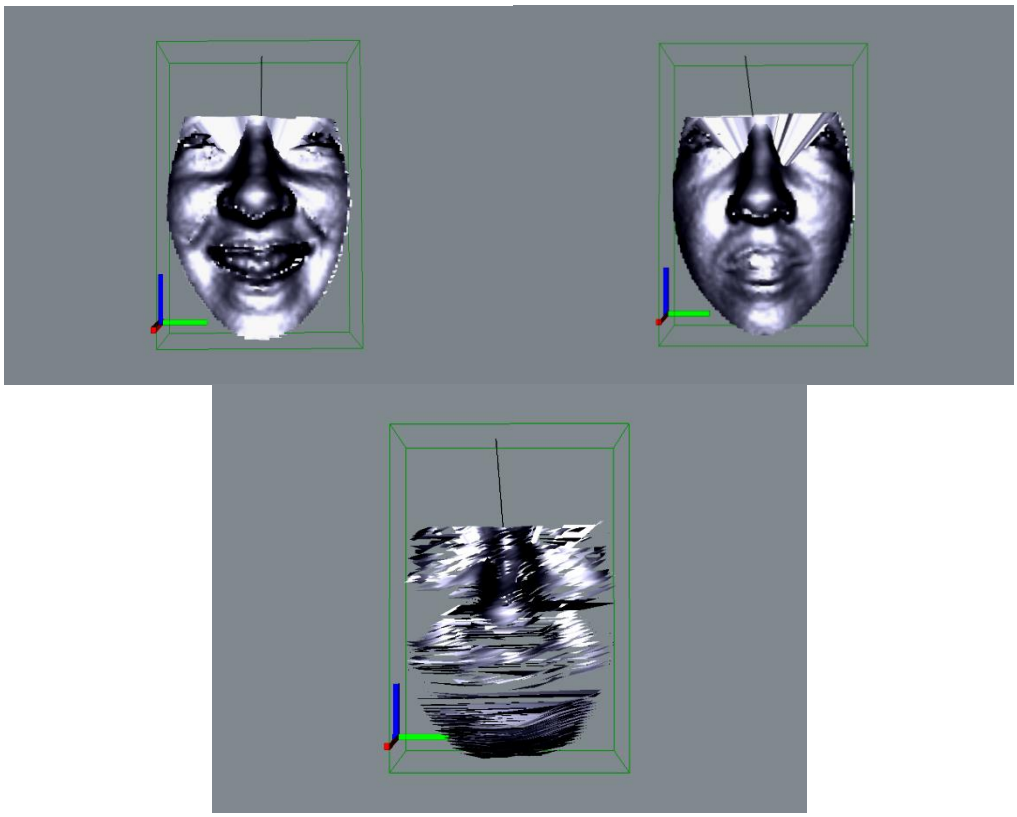
الأول: انقاص نسبة الاستبعاد أي انقاص سرعة عملية التعرف.

الثاني: حساب تمثيل الدوائر المتمركزة من المناطق المقتطعة الأقل تأثر بالتعابير من الوجوه المستخدمة في بناء قاعدة البيانات للحصول على قاعدة بيانات من هذا التمثيل لهذه الوجوه وهذا هو الحل الصحيح وباستخدام هذا الأسلوب يتحسن معدل التعرف على الوجوه بوجود تعابير إلى RR=٨٧,٥٪ من أجل نفس نسبة الاستبعاد ٧١,٨٧٥%.

٣-٣-١ التجزيء غير المنتظم للوجه :non uniform face segmentation

الوجوه التي فيها تعابير يتم تقسيمها إلى مجموعتي اختبار و تدريب. المجموعة التدريبية تستخدم أثناء المعالجة الأولية offline لتحديد أتوماتيكياً مناطق الوجه الأقل تأثراً بالتعابير. نستخدم وجهين لكل شخص في قاعدة البيانات أحدهما بدون وجود تعابير والآخر بوجود تعابير ونقيس الاختلاف بينهما (الفرق بين قيم z للنقاط المتقابلة في الوجهين) والمناطق من الوجوه في قاعدة البيانات التي يكون الاختلاف فيها أقل من عتبة معينة فإنه يتم اقتطاعها لكي يتم استخدامها في عملية التعرف ويتم اختيار العتبة ديناميكياً في كل حالة(لكل شخص) بحيث كلما نقصت العتبة زادت الدقة في التعرف على الوجوه بوجود تعابير والعكس بالعكس(طبعاً يجب الأخذ بعين الاعتبار عند اختيار العتبة المحافظة على معلومات كافية من كل وجه من أجل عملية التعرف) فنختار عتبة تقسيم تستبعد ٣٥% من نقاط الوجه ثلاثي الأبعاد.

الشكل التالي يوضح مثال عن وجه ثلاثي الأبعاد بدون ومع تعابير والقناع المشتق من أجل اقتطاع مناطق الوجه غير الحساسة للتعابير التي فيها تغيرات أقل بوجود تعابير.



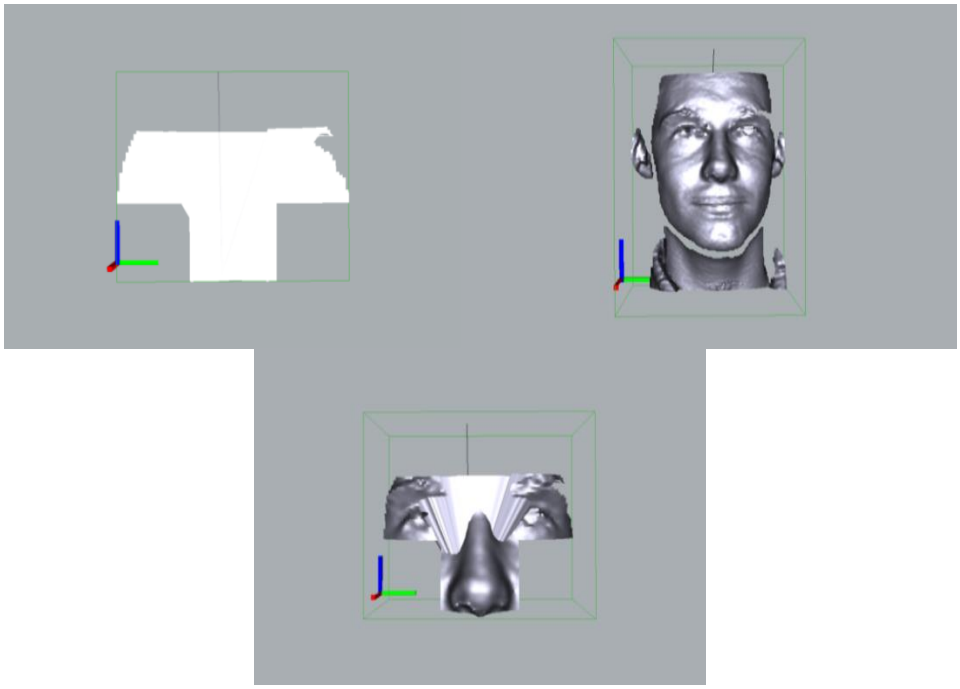
الشكل(٣-٣):مثال عن وجه ثلاثي الأبعاد بدون ومع تعابير. وقناع المناطق غير الحساسة للتعابير

خلال عملية التعرف على شخص أولاً نعالج صورة الوجه الاختباري ثلاثية الأبعاد معالجة أولية كما هو موضح أعلاه ثم نحسب تمثيل الدوائر المتمركزة له ومقارنته مع التمثيل المقابل للوجوه في قاعدة البيانات لاستبعاد الوجوه غير المحتملة وينتج عن عملية المقارنة شعاع قيم خطأ(تعبر عن مقدار التشابه) ذو حجم M (حيث M حجم قاعدة البيانات).

نستبعد الوجوه في قاعدة البيانات التي تكون قيمة الخطأ من أجلها أكبر من غيرها أي ذات الشبه الأقل مع الوجه تحت الاختبار بحيث نحقق نسبة الاستبعاد المناسبة واختيار هذه النسبة عبارة عن عملية مقايضة بين الدقة وكفاءة الاستبعاد (أي بين الدقة والسرعة) و نستخدم الوجوه الباقية في قاعدة البيانات للمقارنة مع الوجه تحت الاختبار بالاستفادة من خوارزمية "النقطة الأقرب التكرارية"(ICP) Iterative Closest Point. ونستخدم المناطق من الوجوه في قاعدة البيانات التي تكون غير حساسة تجاه التعابير من أجل المقارنة لتجنب تأثيرات التعابير لكن ليست هناك حاجة لتجزيء الوجه تحت الاختبار. خطأ التقابل الوسطي بين الوجه تحت الاختبار وأحد وجوه قاعدة البيانات يستخدم كمحصلة التشابه(يعبر عن مقدار التشابه) بينهما(كلما قل الخطأ كان التطابق أكبر).

٣-٣-٢ التجزيء المنتظم للوجه: uniform face segmentation

لا تحتاج هذه الطريقة إلى صور تدريبية طالما أنه يتم اقتطاع الملامح أو المناطق نفسها بشكل منسجم من كل الوجوه (الوجوه في قاعدة البيانات ووجوه الاختبار). بما أن الأنف والجبهة والمنطقة حول العينين هي الأقل حساسية تجاه تعابير الوجه فإننا نقوم باقتطاع هذه المناطق واستخدامها أثناء عملية التعرف. ونقتطع هذه المناطق أتوماتيكياً بواسطة تحديد قناع يقتطع من الوجه: الأنف والعينين ومنطقة الجبهة كما هو موضح بالشكل (٣-٤). ونلاحظ أنه في هذه الحالة ليست هناك حاجة لاستخدام صور تدريبية تحوي تعابير ولذلك نقتطع هذه المناطق من وجه الاختبار قبل عملية المقارنة(فحص التطابق).



الشكل (٣-٤) في اليسار :وجه ثلاثي الأبعاد. في اليمين : قناع. في الأسفل : الوجه ثلاثي الأبعاد بعد عملية الاقتطاع بواسطة القناع .

نقتطع هذه المناطق من الوجوه في قاعدة البيانات أثناء عملية المعالجة (offline). وخلال مرحلة التعرف والتي تكون (online) نقتطع المناطق غير الحساسة تجاه التعابير من الوجه تحت الاختبار ثم

نستبعد جزء من الوجوه في قاعدة البيانات بواسطة مقارنة تمثيل الدوائر المتمركزة للوجوه في قاعدة البيانات مع تمثيل الدوائر المتمركزة للوجه تحت الاختبار كما هو موضح ذلك سابقاً وبعد ذلك تتم مطابقة المناطق غير الحساسة تجاه التعابير من الوجه تحت الاختبار مع المناطق المقابلة للوجوه في قاعدة البيانات (والباقية بعد الاستبعاد) بالاستفادة من خوارزمية ICP. وسوف يرد لاحقاً توضيح نتائج طريقتي تجزيء الوجوه والمقارنة بينهما ويتم اعتماد الطريقة ذات الأداء الأفضل مع تعليل النتائج.

٣-٣-٤ المقارنة (فحص التطابق): Matching

سنوضح الآن كيفية الاستفادة من خوارزمية ICP في عملية المطابقة matching من أجل عملية التعرف التي نقدمها في هذا البحث:

نقوم في البداية بإنشاء تقابلات بين النقاط الأقرب في مجموعتي النقاط ثلاثية الأبعاد ونقوم بتصغير مسافة الخطأ بين هذه النقاط بتطبيق تحويل transformation مؤلف من ازاحة translation ودوران rotation على أحد مجموعتي النقاط.

يتم تكرار هذه العملية حتى يصل خطأ المسافة إلى قيمة اشباع صغرى. ويحتاج هذا الأسلوب إلى عملية مطابقة تقريبية بين مجموعتي النقاط من حيث المكان والوضعية في الفضاء ثلاثي الأبعاد قبل البدء بتنفيذ المطابقة الأدق للمكان والوضعية وذلك لتجنب النهاية المحلية وتستخدم لهذا الغرض خوارزمية تصحيح الوضعية الأتوماتيكية الموضحة أعلاه.

يتم إنشاء التقابلات بين نقاط الوجهين حسب المسافة بين النقطتين وتعتبر هذه النقاط نقاط متقابلة. خطأ المسافة بين مثل هذه النقاط تعطي معلومات أو مؤشر عن التشابه أوعدم التشابه بين الوجهين.

حالما يتم إنشاء التقابلات نحسب متوسط خطأ المسافة ثلاثية الأبعاد بين هذه النقاط المتقابلة ثم نجري دوران وانزياح على أحد الوجهين بهدف تصغير الخطأ ويتم تكرار هذه العملية حتى يصل الخطأ إلى قيمة اشباع أصغرية.

لتكن $P = [x_k,y_k,z_k]^T$ حيث ($k=١.....n_P$) و $G= [x_k,y_k,z_k]^T$ حيث ($k=١.....n_G$) حيث P و G تمثلان نقاط الوجه تحت الاختبار والوجه في قاعدة البيانات على التوالي.

لتكن F تابع يجد النقطة الأقرب في $\hat{}_P$ لكل نقطة من $\hat{}_G$

$$(c , d) = F(\hat{P} , \hat{G})$$

حيث c و d أشعة ذات حجم n_G حيث c_k و d_k تحوي ترتيب النقطة الأقرب في P إلى النقطة k في G ومسافتها عنها و على التوالي أي:

$\forall$ k find $g_k \in \mathbf{G}$ and $p_{ck} \in \mathbf{P} \mid d_k < d_r$ (حيث dr=٣,٥ mm تمثل الحد الأعلى المسموح به للمسافة بين النقطتين لاعتبارهما متقابلتين)

النقطة gi الناتجة تقابل النقطة pi ($\forall i = ١ \ldots N$) حيث N عدد التقابلات بين P و G . خطأ المسافة e الذي سيتم تصغيره يعطى بالعلاقة (٣-٤) ونلاحظ أن e هي خطأ المسافة ثلاثية الأبعاد بين الوجه تحت الاختبار والوجه في قاعدة البيانات. نقوم بتصغير هذا الخطأ بشكل متكرر وتستخدم قيمته النهائية كمحصلة أو قيمة التشابه بين الوجه تحت الاختبار والوجه في قاعدة البيانات.

$$e = ١/N \sum_{i=١}^{N} ||Rg_i + t - p_i|| \quad (٣-٤)$$

لتصغير الخطأ يجب حساب مصفوفة الدوران R وشعاع الانزياح t في العلاقة السابقة وتطبيقهما على أحد الوجهين بحيث تحقق أفضل تطابق في الوضعية والموضع بين الوجهين ويمكن حساب R باستخدام عدةطرق منهاطريقةالتحليل إلى القيم المفردةSingular Value Decomposition (SVD) حيث ستجعل مصفوفة الترابط المتعارض بين pi و gi مصفوفة قطرية مما يعني أفضل تطابق في الوضعية بين الوجهين. وأحد محاسن هذه الطريقة أنه يمكن وبسهولة تعميمها من أجل أي عدد من الأبعاد. مصفوفة الترابط المتعارض K cross-correlation matrix بين pi و gi تعطى بالعلاقة (٣-٥).

$$\mu_p = ١/N \sum_{i=١}^{N} p_i$$

$$\mu_g = ١/N \sum_{i=١}^{N} g_i$$

$$K = ١/N \sum_{i=١}^{N} (g_i - \mu_g)(p_i - \mu_p)^T \quad (٣-٥)$$

$$UAV^T = K \quad (٣-٦)$$

العلاقة (٣-٦) تعبر عن SVD وفيها U,V متعامدة و A قطرية .

التحليل إلى القيم المفردة :Singular Value Decomposition(SVD)

إذا كانت لدينا مصفوفة A أبعادها n*p فإن نظرية SVD تقول أنه بالامكان تحليل المصفوفة A وفق المعادلة التالية :

$$A_{nxp} = U_{nxn} S_{nxp} V^T_{pxp} \quad (٧-٣)$$

حيث :

$$U^T U = I_{nxn}$$
$$V^T V = I_{pxp}$$

أي أن U و V كل منهما مصفوفة متعامدة.

حيث تكون أعمدة U الأشعة المفردة اليسرى (left singular vectors) والمصفوفة S التي لها نفس أبعاد المصفوفة A فيها القيم المفردة singular values وتكون مصفوفة قطرية و V^T صفوفها عبارة عن الأشعة المفردة اليمنى (right singular vectors) .

تمثل SVD تحليلاً(فك) للبيانات الأصلية في نظام احداثيات جديد تكون فيه مصفوفة التغاير قطرية .

لتكن X مصفوفة n*m و X^TX مصفوفة مربعة m*m من المرتبة rank=r ولنعرف ما يلي:

$\{v_1,v_2,.....,v_r\}$ مجموعة من m*١ أشعة خاصة مع القيم الخاصة التي تقابلها $\{\lambda_1,\lambda_2,...,\lambda_r\}$ من أجل المصفوفة X^TX .

$$(\mathbf{X}^T\mathbf{X})\hat{\mathbf{v}}_i = \lambda_i \hat{\mathbf{v}}_i \quad (٨-٣)$$

$\sigma_i \equiv \sqrt{\lambda_i}$ أعداد حقيقية موجبة وتسمى "القيم المفردة" singular values

$\{u_1,u_2,.....,u_r\}$ مجموعة أشعة واحدية متعامدة أبعادها n*١ معرفة بالعلاقة $\hat{\mathbf{u}}_\mathbf{i} \equiv \frac{1}{\sigma_i}\mathbf{X}\hat{\mathbf{v}}_\mathbf{i}$

التعريف الأخير يتضمن خاصتين :

$$\hat{\mathbf{u}}_\mathbf{i} \cdot \hat{\mathbf{u}}_\mathbf{j} = \delta_{ij}$$

$$\|\mathbf{X}\hat{\mathbf{v}}_\mathbf{i}\| = \sigma_i$$

من هذه التعاريف السابقة سنشكل علاقة التحليلdecomposition :

صيغة "القيمة" (value version) لتحليل القيم المفردة هو عبارة فقط عن اعادة صياغة للتعريف الثالث:

$$\mathbf{X}\hat{\mathbf{v}}_i = \sigma_i \hat{\mathbf{u}}_i \qquad (٩-٣)$$

معنى هذه العلاقة أن X مضروبة بشعاع خاص للمصفوفة X^TX مساوي لعدة مرات شعاع آخر u_i. مجموعة الأشعة الخاصة $\{v_١,v_٢,\ldots\ldots,v_r\}$ ومجموعة الأشعة $\{u_١,u_٢,\ldots,u_r\}$ كلتاهما مجموعتي أشعة واحدية متعامدة في فضاء ذو r بعد.

يمكن تلخيص هذه النتيجة من أجل كل الأشعة في عملية ضرب مصفوفية واحدة ونعرف أولاً المصفوفة القطرية ∑ :

$$\mathbf{\Sigma} \equiv \begin{bmatrix} \sigma_{\tilde{1}} & & & & \\ & \ddots & & 0 & \\ & & \sigma_{\tilde{r}} & & \\ & & & 0 & \\ & 0 & & \ddots & \\ & & & & 0 \end{bmatrix}$$

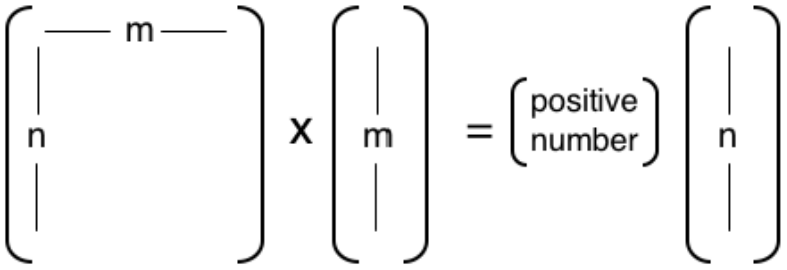
حيث σ١≥σ٢≥.....≥σr تمثل مجموعة القيم المفردة singular values. بشكل مماثل نشكل المصفوفتين U,V التي تضم أشعة واحدة متعامدة :

$$\mathbf{V} = [\hat{\mathbf{v}}_{\tilde{1}}\ \hat{\mathbf{v}}_{\tilde{2}}\ \ldots\ \hat{\mathbf{v}}_{\tilde{m}}]$$
$$\mathbf{U} = [\hat{\mathbf{u}}_{\tilde{1}}\ \hat{\mathbf{u}}_{\tilde{2}}\ \ldots\ \hat{\mathbf{u}}_{\tilde{n}}]$$

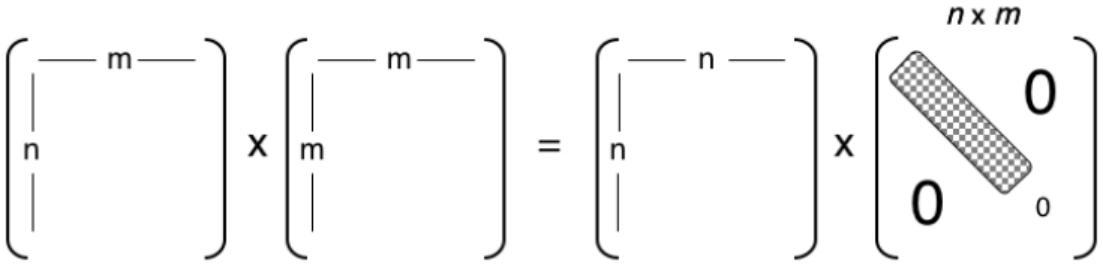
نوضح فيما يلي كيف يتم استخدام المصفوفات السابقة في تشكيل الشكل المصفوفي matrix form لSVD :

$$XV = U\sum \qquad (١٠-٣)$$

الشكل التالي يوضح شكل القيمة value form ل SVD :



الشكل المصفوفي matrix form يوضح بالشكل التالي :



المصفوفتين V و U أبعادهما m*m و n*n على التوالي و ∑ مصفوفة قطرية. بحل هذه العلاقة المصفوفية نحصل على كل المعادلات التي تمثل شكل القيمة value form .

بما أن V مصفوفة متعامدة يمكن ضرب طرفي علاقة الشكل المصفوفي لSVD ب $V^{-1}=V^T$ لنصل إلى الشكل النهائي لعملية التحليل SVD :

$$X = U\sum V^T \quad (١٠-٣)$$

يعتبر هذا التحليل قوياً ومفيداً حيث يقول أن أي مصفوفة X يمكن تحويلها إلى مصفوفة متعامدة ومصفوفة قطرية ومصفوفة متعامدة أخرى(أو دوران rotation وامتداد stretch ودوران آخر rotation).

تفسير SVD :

الشكل النهائي لSVD صعب الفهم ولتسهيل ذلك يجب تفسير شكل القيمة value form لها والذي له الشكل :

$$Xa = kb \quad (١٢-٣)$$

حيث a و b أشعة عمود و k ثابت سلمي. مجموعة الأشعة $\{v_1,v_2,.....,v_m\}$ تقابل a ومجموعة الأشعة $\{u_1,u_2,...,u_n\}$ تقابل b حيث كل من المجموعتين السابقتين تضمان أشعة متعامدة واحدية تشكل فضاءين ذوي أبعاد m و n على التوالي.

$$X = U\sum V^T$$

$$U^TX = \sum V^T$$

$$U^TX = Z \qquad (١٣-٣)$$

حيث نضع أن $Z=\sum V^T$.نلاحظ أن الأعمدة $\{u١,u٢,...,un\}$ في المصفوفة U أصبحت في هذه العلاقة أسطر في المصفوفة U^T وبمقارنة هذه العلاقة مع علاقة PCA نجد أن U^T تمثل مصفوفة دوران تغير تمثيل البيانات من X إلى Z بتغيير المحاور.

أيضاً :

$$XV = \sum U$$

$$(XV)^T = (\sum U)^T$$

$$V^TX^T = U^T\sum$$

$$V^TX^T = Z \qquad (١٤-٣)$$

أيضاً قمنا بوضع $Z = U^T\sum$ ونلاحظ أن أسطر V^T أو أعمدة V تشكل أشعة واحدية متعامدة لتحويل X^T إلى Z (V تمثل مصفوفة دوران).

تعطى مصفوفة الدوران (rotation matrix) R وشعاع الانزياح (translation vector) t بالعلاقتين (٣-١٥) و(٣-١٦) على التوالي :

$$R = VU^T \qquad (١٥-٣)$$

$$t = \mu_p - R\mu_g \qquad (١٦-٣)$$

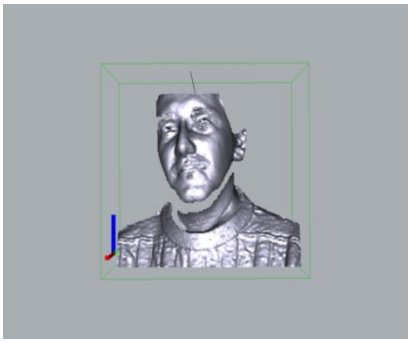
R مسقط قطبي ل K. إذا كان $det(R)=-١$ فإن هذا يعطي أن هناك انعكاساً فتعطى R بالعلاقة:

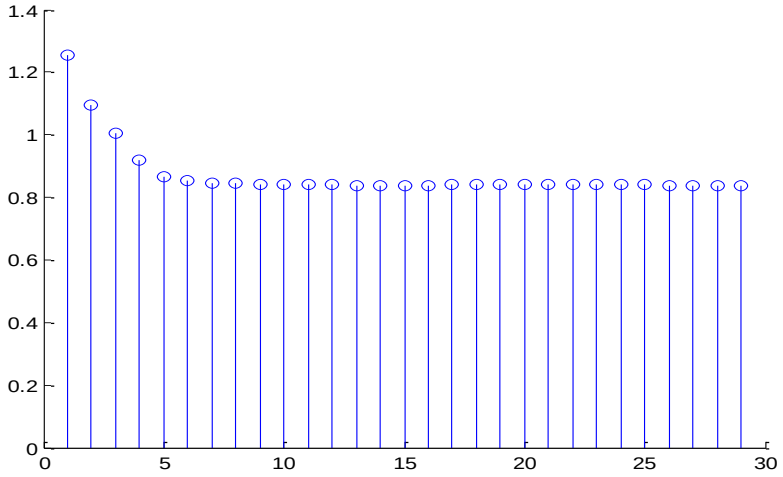
$$R = V \begin{pmatrix} ١ & ٠ & ٠ \\ ٠ & ١ & ٠ \\ ٠ & ٠ & \det(UV^T) \end{pmatrix} U^T \qquad (١٧-٣)$$

إذاً تبدأ عملية التعرف بحساب مسافة الخطأ بين النقاط المتقابلة ثم تطبيق تحويل (ازاحة ودوران على أحد الوجهين) بهدف تصغير الخطأ ثم تحديد من جديد النقاط المتقابلة في مجموعتي النقاط ثلاثية الأبعاد التي تمثل الوجهين وهكذا تكرر العملية أعلاه حتى الوصول إلى قيمة إشباع صغرى للخطأ. عندما تصبح قيمة الخطأ الجديدة قريبة جداً من قيمة الخطأ في المرحلة التي قبلها(أصغر من عتبة معينة dr/١٠٠٠) وعندما يتحقق ذلك عدة مرات تعتبر قيمة الخطأ قد استقرت عند قيمة الإشباع الصغرى لها.

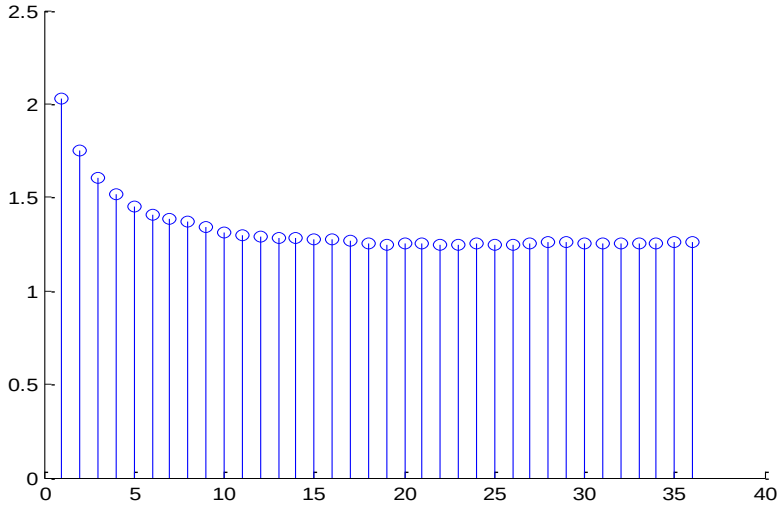
وبعد اجراء هذه العملية بين الوجه تحت الاختبار والوجوه في قاعدة البيانات نحصل على شعاع من قيم الخطأ الذي يعبر عن مقدار التشابه بين الوجه تحت الاختبار والوجوه في قاعدة البيانات بحيث يقابل أقل خطأ أكبر تشابه وبالتالي يتم التعرف على الشخصية المقابلة للوجه تحت الاختبار على أنه الشخصية المقابلة للوجه في قاعدة البيانات صاحب الخطأ الأقل.

الأشكال التالية توضح تغير قيمة الخطأ من أجل الشخصية رقم ١١ الموضحة بالشكل :

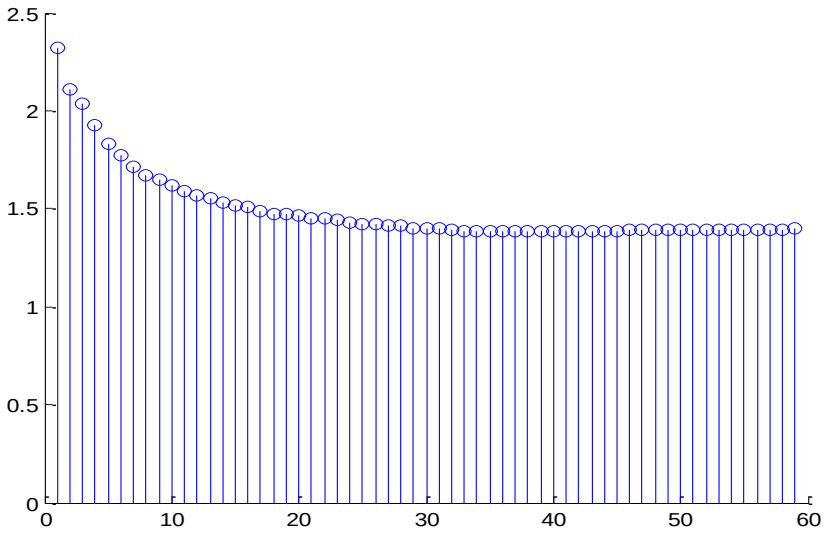




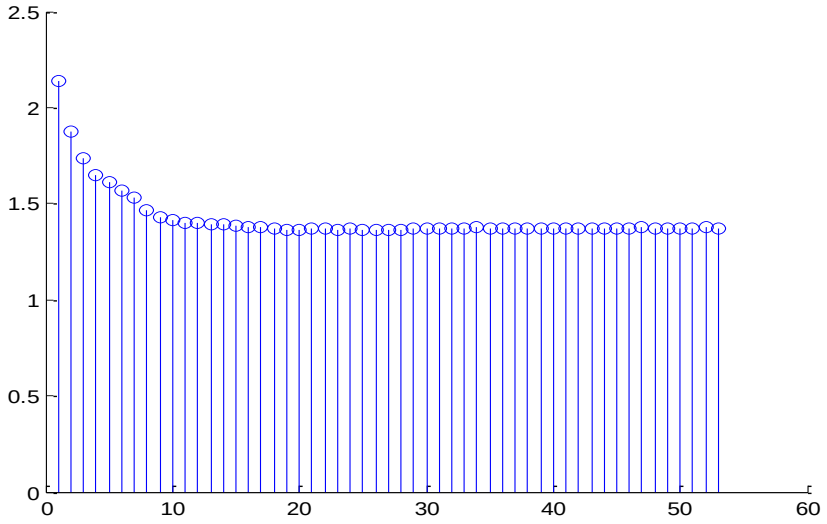
الخطأ من أجل الشخصية رقم ١١ في قاعدة البيانات



الخطأ من أجل الشخصية رقم ٦ في قاعدة البيانات



الخطأ من أجل الشخصية رقم ٨ في قاعدة البيانات



الخطأ من أجل الشخصية رقم ١ في قاعدة البيانات

الشكل(٣-٦) تناقص الخطأ

الفصل الرابع

نتائج عملية التعرف

سوف نوضح في هذا الفصل نتائج عملية التعرف باستخدام الأسلوب الموضح في هذا البحث والذي نفذناه بواسطة MATLAB على ٣٢ وجه من قاعدة البيانات التي ذكرناها في الفصل الأول حيث توضح الجداول في هذا الفصل نتائج عملية التعرف من أجل حالة معينة من أجل الوجوه من ١ إلى ٢٠ في قاعدة البيانات(من أجل التوضيح) ثم في نهاية الفصل نوضح محصلة النتائج وكذلك نبين معدل زمن عملية التعرف الواحدة من أجل حاسبIntel(R)celeron(R)CPU٤٥٠@١,٦٨ GHz.

true في الجداول تشير إلى عملية تعرف ناجحة

false في الجداول تشير إلى عملية تعرف فاشلة أو خاطئة

وكذلك تبين معدل التعرف recognition rate(RR) من أجل حالات مختلفة للوجه تحت الاختبار وحالات عمل مختلفة بالنسبة لطريقة التعامل مع التعابير في الوجوه من أجل بناء قاعدة البيانات.

الحالات المختلفة للوجه تحت الاختبار تتضمن:

- وضعيات مختلفة: أمامية – جانبية – ...

- وجوه بتعابير وبدون تعابير.

حالات العمل المختلفة بالنسبة لطريقة التعامل مع تعابير الوجه من أجل بناء قاعدة البيانات:

-عدم تجزيء للوجوه إلى مناطق حساسة وأقل حساسية لتعابير الوجه.

-حالة التجزيء المنتظم للوجوه.

-حالة التجزيء غير المنتظم للوجوه.

حيث توضح الجداول التالية النتائج(معدل التعرف) من أجل الحالات السابقة وتوضح كذلك علاقة الدقة مع سرعة التعرف عبر مقارنة النتائج من أجل نسب استبعاد مختلفة باستخدام تكمية للتمثيل ثلاثي الأبعاد للوجه.

الجدول التالي يوضح النتائج من أجل الوجوه بوضعية أمامية frontal بدون تعابير من أجل نسبة استبعاد أولية ٧١,٨٧٥% بواسطة استخدام تمثيل الدوائر المتمركزة باستخدام التجزيء غير المنتظم للوجوه مع عتبة تقسيم تستبعد ٣٥% من نقاط الوجه ثلاثي الأبعاد:

رقم الشخصية	وضعية أمامية ١ بدون تعابير	وضعية أمامية ٢ بدون تعابير
١	true	true
٢	true	true
٣	true	true
٤	true	true
٥	true	true
٦	true	true
٧	true	true
٨	true	true
٩	true	true
١٠	true	true
١١	true	true
١٢	true	true
١٣	true	true
١٤	true	true
١٥	true	true
١٦	true	true
١٧	true	true
١٨	true	true
١٩	true	true
٢٠	true	true

الجدول(٤-١)

معدل التعرف (RR=١٠٠%) من أجل الوجوه بوضعية أمامية frontal بدون تعابير من أجل نسبة استبعاد أولية ٧١,٨٧٥% بواسطة استخدام تمثيل الدوائر المتمركزة.

الجدول التالي يوضح النتائج من أجل الوجوه بوضعية أمامية frontal بدون تعابير من أجل نسبة استبعاد أولية ٧١,٨٧٥% بواسطة استخدام تمثيل الدوائر المتمركزة وباستخدام التجزيء المنتظم الموحد للوجوه:

رقم الشخصية	وضعية أمامية ١ بدون تعابير	وضعية أمامية ٢ بدون تعابير
١	true	true
٢	true	true
٣	true	true
٤	true	true
٥	true	true
٦	true	true
٧	true	true
٨	true	false
٩	true	true
١٠	true	true
١١	true	true
١٢	true	true
١٣	true	true
١٤	true	true
١٥	true	true
١٦	true	true
١٧	true	true
١٨	true	true
١٩	true	true
٢٠	true	true

الجدول(٤-٢)

معدل التعرف (RR=%٩٦,٨٨) من أجل الوجوه بوضعية أمامية frontal بدون تعابير من أجل نسبة استبعاد أولية ٧١,٨٧٥% بواسطة استخدام تمثيل الدوائر المتمركزة.

الجدول التالي يوضح النتائج من أجل الوجوه بوضعية أمامية frontal بدون تعابير من أجل نسبة استبعاد أولية ٧١,٨٧٥% بواسطة استخدام تمثيل الدوائر المتمركزة وبدون تجزيء للوجوه أي بدون اقتطاع المناطق الأقل حساسية للتعابير:

رقم الشخصية	وضعية أمامية ١ بدون تعابير	وضعية أمامية ٢ بدون تعابير
١	true	true
٢	true	true
٣	true	true
٤	true	true
٥	true	false(true for rank=٣)
٦	true	true
٧	true	true
٨	true	false(true for rank=٨)
٩	true	true
١٠	true	true
١١	true	true
١٢	true	true
١٣	true	true
١٤	true	true
١٥	true	true
١٦	true	true
١٧	true	true
١٨	true	true
١٩	true	true
٢٠	true	true

الجدول(٤-٣)

معدل التعرف (RR=%٩٥) من أجل الوجوه بوضعية أمامية frontal بدون تعابير من أجل نسبة استبعاد أولية ٧١,٨٧٥% بواسطة استخدام تمثيل الدوائر المتمركزة.

الجدول التالي يوضح النتائج من أجل الوجوه بوضعيات مختلفة بدون تعابير من أجل نسبة استبعاد أولية ٧١,٨٧٥% بواسطة استخدام تمثيل الدوائر المتمركزة وباستخدام تجزيء غير منتظم للوجوه مع عتبة تقسيم تستبعد ٣٥% من نقاط الوجه ثلاثي الأبعاد:

أولاً : الدوران حول المحور y :

رقم الشخصية	وضعيةجانبية١(دوران حول المحور y نحو اليسار أي عكس عقارب الساعة)	وضعيةجانبية٢(دوران حول المحور y نحو اليمين أي مع عقارب الساعة)	وضعية جانبية دوران حول المحور y ٩٠ درجة(تقريباً)نحو اليسار	وضعية جانبية دوران حول المحور y ٩٠ درجة(تقريباً) نحو اليمين
١	true	true	false	true
٢	true	true	false	false
٣	true	true	false	false
٤	true	true	false	false
٥	false(true for rank=٧)	true	false	false
٦	true	true	false	false
٧	true	true	false	false
٨	true	false(true for rank=٢١)	false	false
٩	true	true	false	false
١٠	true	true	true	false
١١	true	true	false	false
١٢	true	true	false	false
١٣	true	true	true	false
١٤	true	true	false	false
١٥	true	true	false	false
١٦	true	true	false	false
١٧	true	true	false	false
١٨	true	true	false	false
١٩	true	true	false	true

٢٠	true	true	false	false

الجدول(٤-٤)

معدل التعرف (RR=%٩٥,٣١) من أجل الوجوه بوضعيات مختلفة (دوران حول المحور y أقل من ٤٥ درجة) من أجل نسبة استبعاد أولية ٧١,٨٧٥% بواسطة استخدام تمثيل الدوائر المتمركزة

معدل التعرف(RR=%١٤,٠٦) من أجل الوجوه بوضعيات مختلفة (دوران حول المحور y أكبر من ٤٥ درجة) من أجل نسبة استبعاد أولية ٧١,٨٧٥% بواسطة استخدام تمثيل الدوائر المتمركزة

ثانياً : الدوران حول المحور x :

رقم الشخصية	وضعية جانبية١ (دوران حول المحورxنحو الأعلى)	وضعية جانبية٢(دوران حول المحورxنحو الأسفل)
١	true	true
٢	true	true
٣	true	true
٤	true	true
٥	true	true
٦	true	true
٧	true	true
٨	true	false
٩	true	true
١٠	true	true
١١	true	true
١٢	true	true
١٣	true	true
١٤	true	false(wrong nose detection)
١٥	true	true
١٦	true	true
١٧	fasle(wrong nose detection)	true
١٨	fasle(wrong nose detection)	true

١٩	fasle(wrong nose detection)	true
٢٠	true	true

الجدول(٤-٥)

معدل التعرف(RR=%٨٧,٥) من أجل الوجوه بوضعيات مختلفة (دوران حول المحورx حتى ٣٥ درجة) من أجل نسبة استبعاد أولية ٧١,٨٧٥% بواسطة استخدام تمثيل الدوائر المتمركزة والأخطاء في هذه الحالة في معظمها ناتجة عن تحديد خاطئ لقمة الأنف وبالتالي اقتطاع منطقة خاطئة والنتيجة عملية تعرف فاشلة.

الجدول التالي يوضح النتائج من أجل الوجوه بوضعيات مختلفة بدون تعابير من أجل نسبة استبعاد أولية ٧١,٨٧٥% بواسطة استخدام تمثيل الدوائر المتمركزة وباستخدام تجزيء منتظم موحد للوجوه :

أولاً : الدوران حول المحور y :

رقم الشخصية	دوران حول المحور yنحواليسار	دوران حول المحور yنحواليمين	وضعيةجانبيةprofile دوران نحواليسار	وضعيةجانبيةprofile دوران نحواليمين
١	true	true	false	false
٢	true	true	false	false
٣	true	true	false	false
٤	true	true	false	false
٥	true	true	false	false
٦	true	true	false	false
٧	true	true	false	false
٨	true	false(true for rank=٢٠)	false	false
٩	true	true	false	false
١٠	true	true	true	false
١١	true	true	false	false
١٢	true	true	false	false
١٣	true	true	true	false
١٤	false(true for rank=٤)	true	false	false
١٥	true	true	false	false
١٦	true	true	false	false

١٧	true	true	false	false
١٨	true	true	false	false
١٩	true	true	false	true
٢٠	true	true	false	false

الجدول(٤-٦)

معدل التعرف (RR=%٩٢,١٩) من أجل الوجوه بوضعيات مختلفة (دوران حول المحور y أقل من ٤٥ درجة) من أجل نسبة استبعاد أولية ٧١,٨٧٥% بواسطة استخدام تمثيل الدوائر المتمركزة.

معدل التعرف (RR=%١٠.٩٤) من أجل الوجوه بوضعيات مختلفة (دوران حول المحور y أكبر من ٤٥ درجة) من أجل نسبة استبعاد أولية ٧١,٨٧٥% بواسطة استخدام تمثيل الدوائر المتمركزة.

ثانياً : الدوران حول المحور x :

رقم الشخصية	دوران حول المحورxنحو الأعلى	دوران حول المحورxنحو الأسفل
١	true	true
٢	true	true
٣	true	true
٤	true	true
٥	true	true
٦	true	true
٧	true	true
٨	false	false(true for rank=٤)
٩	true	true
١٠	true	true
١١	true	false
١٢	true	true
١٣	true	true
١٤	true	false(wrong nose detection
١٥	true	true
١٦	true	true

١٧	fasle(wrong nose detection)	true
١٨	fasle(wrong nose detection)	true
١٩	fasle(wrong nose detection)	false
٢٠	true	false

الجدول(٤-٧)

معدل التعرف (RR=%٧٥) من أجل الوجوه بوضعيات مختلفة (دوران حول المحورx حتى ٣٥ درجة) من أجل نسبة استبعاد أولية ٧١,٨٧٥% بواسطة استخدام تمثيل الدوائر المتمركزة. الأخطاء في هذه الحالة في معظمها ناتجة عن تحديد خاطئ لقمة الأنف وبالتالي اقتطاع منطقة خاطئة والنتيجة عملية تعرف فاشلة.

الجدول التالي يوضح النتائج من أجل الوجوه بوضعيات مختلفة بدون تعابير من أجل نسبة استبعاد أولية ٧١,٨٧٥% بواسطة استخدام تمثيل الدوائر المتمركزة وبدون تجزيء للوجوه :

أولاً : الدوران حول المحور y :

رقم الشخصية	دوران حول المحور yنحواليسار	دوران حول المحور yنحواليمين	وضعيةجانبيةprofile دوران نحواليسار	وضعيةجانبيةprofile دوران نحواليمين
١	true	true	false	false
٢	true	true	false	false
٣	true	true	false	false
٤	true	true	false	false
٥	true	true	false	false
٦	true	true	false	false
٧	false	true	false	false
٨	false	false(true for rank=١٠)	false	false
٩	true	true	false	false
١٠	true	true	true	false
١١	true	true	false	false
١٢	true	true	false	false
١٣	true	true	false	false
١٤	true	true	false	false

١٥	true	true	false	false
١٦	true	true	false	false
١٧	true	true	false	false
١٨	true	true	false	false
١٩	true	true	false	true
٢٠	true	true	false	false

الجدول(٤-٨)

معدل التعرف (RR=%٩٢,٥) من أجل الوجوه بوضعيات مختلفة (دوران حول المحور y أقل من ٤٥ درجة) من أجل نسبة استبعاد أولية ٧١,٨٧٥% بواسطة استخدام تمثيل الدوائر المتمركزة

معدل التعرف (RR=%٥) من أجل الوجوه بوضعيات مختلفة (دوران حول المحور y أكبر من ٤٥ درجة) من أجل نسبة استبعاد أولية ٧١,٨٧٥% بواسطة استخدام تمثيل الدوائر المتمركزة

ثانياً : الدوران حول المحور x :

رقم الشخصية	دوران حول المحور xنحو الأعلى	دوران حول المحور xنحو الأسفل
١	true	true
٢	true	true
٣	true	true
٤	true	true
٥	true	true
٦	true	true
٧	true	false
٨	false	true
٩	true	true
١٠	true	true
١١	true	true
١٢	false	true
١٣	false(true for rank=٤)	false(true for rank=٣)
١٤	true	false(wrong nose detection)

١٥	true	true
١٦	true	true
١٧	fasle(wrong nose detection)	true
١٨	fasle(wrong nose detection)	true
١٩	fasle(wrong nose detection)	true
٢٠	true	fasle(wrong nose detection)

الجدول(٤-٩)

معدل التعرف (RR=%٧١,٤٣) من أجل الوجوه بوضعيات مختلفة (دوران حول المحورx حتى ٣٥ درجة) من أجل نسبة استبعاد أولية ٧١,٨٧٥%بواسطة استخدام تمثيل الدوائر المتمركزة. الأخطاء في هذه الحالة في معظمها ناتجة عن تحديد خاطئ لقمة الأنف وبالتالي اقتطاع منطقة خاطئة والنتيجة عملية تعرف فاشلة.

النتائج من أجل التعرف على الوجوه ثنائية الأبعاد باستخدام الوجوه الخاصة eigenfaces :

رقم الشخصية	وضعية أمامية	وضعيةجانبية(دوران حول المحورy)	وضعية دوران حول المحور x
١	true	false	true
٢	true	true	true
٣	true	true	true
٤	true	false	false
٥	true	true	false
٦	true	true	true
٧	true	true	true
٨	true	true	true
٩	false	false	true
١٠	true	true	true
١١	true	false	false
١٢	true	true	true
١٣	true	false	false
١٤	true	true	true

١٥	true	false	false
١٦	true	false	false
١٧	false	false	false
١٨	true	false	false
١٩	true	false	false
٢٠	false	false	false

الجدول(٤-١٠)

معدل التعرف على الوجوه يساوي ٨٥% باستخدام طريقة الوجوه الخاصة (ثنائية الأبعاد) من أجل وضعية أمامية frontal pose

ومعدل التعرف ٤٧,٥% من أجل وجوه بوضعية غير الوضعية الأمامية (أي هناك دوران حول أحد المحاور).

ومع تغير الوضعية (أمامية-جانبية-...) يصبح معدل التعرف بشكل عام ٦٠%.

ومنه نلاحظ التحسن الكبير في معدل التعرف في خوارزمية التعرف على الوجوه ثلاثية الأبعاد عند تغير الوضعية pose variation مقارنة مع طريقة الوجوه الخاصة بالنسبة للوجوه ثنائية الأبعاد والسبب هو تصحيح الوضعية pose correction للوجوه ثلاثية الأبعاد الذي تم القيام به في هذا البحث ثم القيام بعملية التعرف.

النتائج من أجل وجوه فيها تعابير من أجل نسبة استبعاد ٠% (أي بدون استبعاد)مع تجزيء غير منتظم للوجوه بعتبة تقسيم تستبعد ٣٥% من نقاط الوجوه في قاعدة البيانات:

رقم الشخصية	وضعية أمامية مع تعابير (ابتسامة)	وضعية أمامية مع تعابير (ضحكة)	تعبير عشوائي
١	true	true	false(true for rank=٣)
٢	true	true	false(true for rank=٣)
٣	true	true	false(true for rank=٤)
٤	true	true	true
٥	true	true	true
٦	true	true	true
٧	true	true	false(true for rank=٢)
٨	true	true	false(true for rank=٢)
٩	true	true	true
١٠	true	true	false(true for rank=٢)
١١	true	true	true
١٢	true	true	true
١٣	true	true	true
١٤	true	true	false
١٥	true	true	true
١٦	true	true	true

١٧	true	true	true
١٨	true	true	true
١٩	true	true	true
٢٠	true	true	true

الجدول(٤-١١)

معدل التعرف يساوي٨٩,٥٨%=RRمن أجل الوجوه بوجود تعابير من أجل نسبة استبعاد٠%(أي بدون استبعاد).وينخفض معدل التعرف عند استخدام عتبة تجزيءأكبرمن المستخدمةفي الجدول أعلاه والسبب هو تقليل التشدد في اقتطاع المناطق الأقل حساسية للتعابير من الوجوه المستخدمة في بناء قاعدة البيانات مما يجعل معدل التعرف يتأثر سلباً أكثر بوجود التعابير في الوجوه تحت الاختبار.

النتائج من أجل وجوه فيها تعابير من أجل نسبة استبعاد ٧١,٨٧٥% مع عتبة تقسيم تستبعد ٣٥% من نقاط الوجه ثلاثي الأبعاد:

رقم الشخصية	وضعية أمامية مع تعابير (ابتسامة)	وضعية أمامية مع تعابير (ضحكة)	تعبير عشوائي
١	true	true	false
٢	true	true	false
٣	true	true	false
٤	true	false	true
٥	true	false	true
٦	true	true	true
٧	true	true	false
٨	true	true	false
٩	true	true	true
١٠	true	true	false
١١	true	true	true
١٢	true	true	true
١٣	true	true	true
١٤	true	true	false(بسبب الشعر الذي

			يغطي جزء من الوجه مما أدى إلى تحديد خاطئ لقمة الأنف)
١٥	true	true	true
١٦	true	true	true
١٧	true	true	true
١٨	true	true	true
١٩	true	true	true
٢٠	true	true	true

الجدول(٤-١٢)

معدل التعرف يساوي(RR=%٨٦,٤٦) من أجل الوجوه بوجود تعابير في حالة التجزيء غير المنتظم للوجوه من أجل نسبة استبعاد ٧١,٨٧٥%.

النتائج من أجل وجوه فيها تعابير من أجل نسبة استبعاد ٠% (أي بدون استبعاد) وباستخدام تجزيء منتظم موحد للوجوه (منطقة محددة موحدة لكل الوجوه):

رقم الشخصية	وضعية أمامية مع تعابير (ابتسامة)	وضعية أمامية مع تعابير (ضحكة)	تعبير عشوائي
١	true	true	true
٢	true	true	true
٣	true	true	true
٤	true	false(true for rank=٣)	true
٥	true	false(true for rank=٢)	true
٦	true	true	true
٧	true	true	true
٨	true	true	true
٩	true	false(true for rank=٤)	true
١٠	true	true	true
١١	true	true	true

١٢	true	true	true
١٣	true	true	true
١٤	true	true	false(بسبب الشعر الذي يغطي جزء من الوجه مما أدى إلى تحديد خاطئ لقمة الأنف)
١٥	true	true	true
١٦	true	true	true
١٧	true	true	true
١٨	true	true	true
١٩	true	true	true
٢٠	true	true	true

الجدول(٤-١٣)

معدل التعرف ٩٣,٧٥%=RR من أجل الوجوه بوجود تعابير وباستخدام تجزيء منتظم موحد للوجوه.

النتائج من أجل وجوه فيها تعابير من أجل نسبة استبعاد ٧١,٨٧٥% وباستخدام تجزيء منتظم موحد للوجوه (منطقة محددة موحدة لكل الوجوه):

رقم الشخصية	وضعية أمامية مع تعابير (ابتسامة)	وضعية أمامية مع تعابير (ضحكة)	تعبير عشوائي
١	true	true	true
٢	true	true	true
٣	true	true	true
٤	true	false	true
٥	true	false	true
٦	true	true	true
٧	true	true	false
٨	false	true	false
٩	true	false	true
١٠	true	true	true
١١	true	true	true

١٢	true	true	true
١٣	false	true	true
١٤	true	true	false(بسبب الشعر الذي يغطي جزء من الوجه مما أدى إلى تحديد خاطئ لقمة الأنف)
١٥	true	true	true
١٦	true	true	true
١٧	true	true	true
١٨	true	true	true
١٩	true	true	true
٢٠	true	true	true

الجدول(٤-١٤)

معدل التعرف يساوي ٨٧,٥% من أجل الوجوه بوجود تعابير وباستخدام تجزيء منتظم موحد للوجوه من أجل نسبة استبعاد ٧١,٨٧٥%.

النتائج من أجل وجوه فيها تعابير من أجل نسبة استبعاد ٨١,٢٥% وباستخدام تجزيء منتظم موحد للوجوه (منطقة محددة موحدة لكل الوجوه):

رقم الشخصية	وضعية أمامية مع تعابير (ابتسامة)	وضعية أمامية مع تعابير (ضحكة)	تعبير عشوائي
١	true	true	true
٢	true	true	true
٣	true	true	true
٤	true	false	true
٥	true	false	true
٦	true	true	true
٧	true	true	false
٨	false	true	false
٩	true	false	true
١٠	true	true	true

١١	true	true	true
١٢	true	true	true
١٣	false	true	true
١٤	true	true	false
١٥	true	true	true
١٦	true	true	true
١٧	true	true	true
١٨	true	true	true
١٩	false	false	true
٢٠	true	true	true

الجدول(٤-١٥)

معدل التعرف ٨١,٢٥%=RR من أجل الوجوه بوجود تعابير وباستخدام تجزيء منتظم موحد للوجوه من أجل نسبة استبعاد ٨١,٢٥%.

نوضح فيما يلي محصلة النتائج:

- معدل التعرف على الوجوه مع تعابير بدون استخدام تمثيل الدوائر المتمركزة بدون تجزيء للوجوه هي ٦٣,٥٤%.

- معدل التعرف على الوجوه بدون تعابير ومع تعابير بدون استخدام تمثيل الدوائر المتمركزة من أجل تجزيء غير منتظم للوجوه مع عتبة تقسيم تستبعد ٣٥% من نقاط الوجه ثلاثي الأبعاد هي ٩٥,٣١% و ٨٩,٥٨% على التوالي ومتوسط زمن عملية التعرف ١٥٦,٥٤ ثانية.

- معدل التعرف على الوجوه بدون تعابير ومع تعابير باستخدام تمثيل الدوائر المتمركزة مع نسبة استبعاد ٧١,٨٧٥% من أجل تجزيء غير منتظم للوجوه مع عتبة تقسيم تستبعد ٣٥% من نقاط الوجه ثلاثي الأبعاد هي ٩٤,٢٧% و ٨٦,٤٦% على التوالي ومتوسط زمن عملية التعرف ٥٥,٤٧ ثانية.

– معدل التعرف على الوجوه بدون تعابير ومع تعابير بدون استخدام تمثيل الدوائر المتمركزة من أجل تجزيء منتظم للوجوه هي ٩٤,٢٧% و ٩٣,٧٥% على التوالي.

– معدل التعرف على الوجوه بدون تعابير ومع تعابير باستخدام تمثيل الدوائر المتمركزة مع نسبة استبعاد ٧١,٨٧٥% من أجل تجزيء منتظم للوجوه هي ٨٦,٩٨% و ٨٧,٥% على التوالي ومتوسط الزمن لعملية التعرف ٢٧,٥ ثانية وينخفض معدل التعرف على الوجوه بوجود تعابير إلى ٨١,٢٥% عند حساب تمثيل الدوائر المتمركزة للوجه قبل اقتطاع المناطق الأقل حساسية.

– معدل التعرف على الوجوه بدون تعابير ومع تعابير باستخدام تمثيل الدوائر المتمركزة مع نسبة استبعاد ٨١,٢٥% من أجل تجزيء منتظم للوجوه هي ٨١,٢٥% و ٨١,٢٥% على التوالي ومتوسط زمن عملية التعرف ٢٢,٧٨ ثانية.

نلاحظ أن الخوارزمية تتأثر سلباً بوجود التعابير وينخفض معدل التعرف من أجل الوجوه بوجود تعابير لأن الوجوه المستخدمة في قاعدة البيانات وجوه ليس فيها تعابير لذلك عند عدم اقتطاع المناطق غير الحساسة للتعابير من الوجوه المستخدمة في بناء قاعدة البيانات ينخفض معدل التعرف إلى ٦٣,٥٤% عند محاولة التعرف على وجوه بوجود تعابير لذلك قمنا في هذا البحث باقتطاع مناطق الوجه الأقل تأثراً بالتعابير واستخدامها في قاعدة البيانات من أجل عملية التعرف مما أدى إلى تحسن معدل التعرف إلى ٨٩,٥٨% في حالة التجزيء أو الاقتطاع غير المنتظم وإلى ٩٣,٧٥% في حالة التجزيء أو الاقتطاع المنتظم مما يوضح أيضاً أن التجزيء المنتظم يعطي نتائج أفضل من ناحية الدقة.

هناك مقايضة بين نسبة الكفاءة لعملية الاستبعاد أي السرعة ودقة عملية التعرف فكلما زادت نسبة الكفاءة (زادت السرعة) نقصت دقة التعرف والعكس بالعكس والسبب في ذلك هو تمثيل الدوائر المتمركزة للوجه ثلاثي الأبعاد وهو تمثيل غير كامل يعطي نتائج تطابق تقريبية ويتضح ذلك من مقارنة زمن

عملية التعرف ومعدل التعرف من أجل الوجوه بوجود تعابير في حالة التجزيء المنتظم للوجوه والموضحين في النتائج أعلاه من أجل نسبتي استبعاد ٧١,٨٧٥% و ٨١,٢٥% حيث نلاحظ مع زيادة نسبة الاستبعاد انخفاض دقة التعرف من ٨٧,٥% إلى ٨١,٢٥% وزيادة سرعة التعرف (زمن عملية التعرف ينخفض من ٢٧,٥ إلى ٢٢,٧٨ ثانية) وهذا يبين وجود المقايضة بين دقة وسرعة التعرف.

إن أسلوب التجزيء المنتظم أسرع في عملية التعرف بسبب حجم البيانات(عدد النقاط) الأقل التي تمثل المناطق الأقل حساسية تجاه التعابير للوجوه في قاعدة البيانات(استخدام مناطق أقل من الوجه في حالة التجزيء المنتظم للوجوه منه في حالة التجزيء غير المنتظم للوجوه) ويتضح ذلك من مقارنة متوسط زمن عملية التعرف لكل أسلوب من أجل نسبة الاستبعاد ٧١,٨٧٥% والموضح بالنتائج أعلاه. إلا أن ذلك يقلل من دقة تمثيل الدوائر المتمركزة للمناطق الباقية بعد الاقتطاع مما يجعل الدقة في طريقة التجزيء المنتظم أكثر تأثراً مع زيادة نسبة الاستبعاد.

لذلك سيعتمد هذا البحث طريقة التجزيء المنتظم للوجوه مع النتائج:

من أجل نسبة استبعاد ٠% مع معدل تعرف ٩٤,٢٧% و ٩٣,٧٥% من أجل الوجوه بدون ومع تعابير على التوالي(كما يبين ذلك النتائج أعلاه) ومع سرعة تعرف أكبر من حالة استخدام التجزيء غير المنتظم للوجوه.

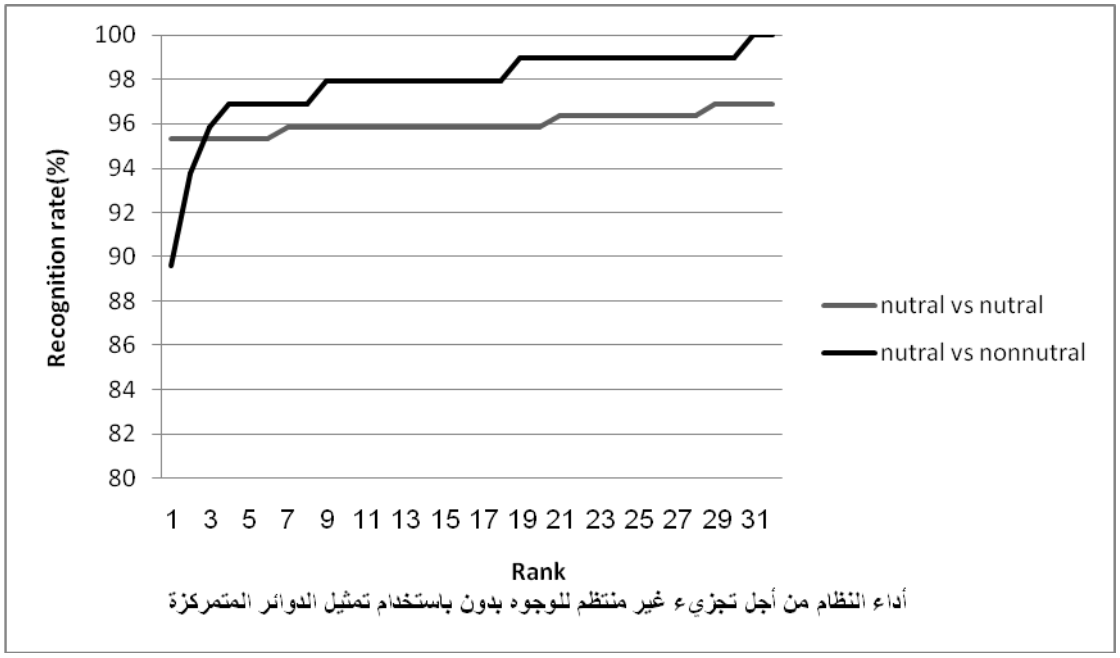
من أجل نسبة استبعاد ٧١,٨٧٥% مع معدل تعرف ٨٦,٩٨% و ٨٧,٥% من أجل الوجوه بدون ومع تعابير على التوالي(كما يبين ذلك النتائج أعلاه) ومع سرعة تعرف أكبر من حالة استخدام التجزيء غير المنتظم للوجوه.

وبمقارنة النتائج في الجدول الذي يوضح نتائج تقنية PCA والوارد في الفصل الأول مع نتائج طريقتنا في هذا البحث للتعرف على الوجوه ثلاثية الأبعاد(بدون تعابير وبدون أسلوب التسريع لعملية التعرف المتبع في هذا البحث أي بنسبة استبعاد ٠%) نجد عند مقارنة معدل التعرف RR أن طريقتنا حققت

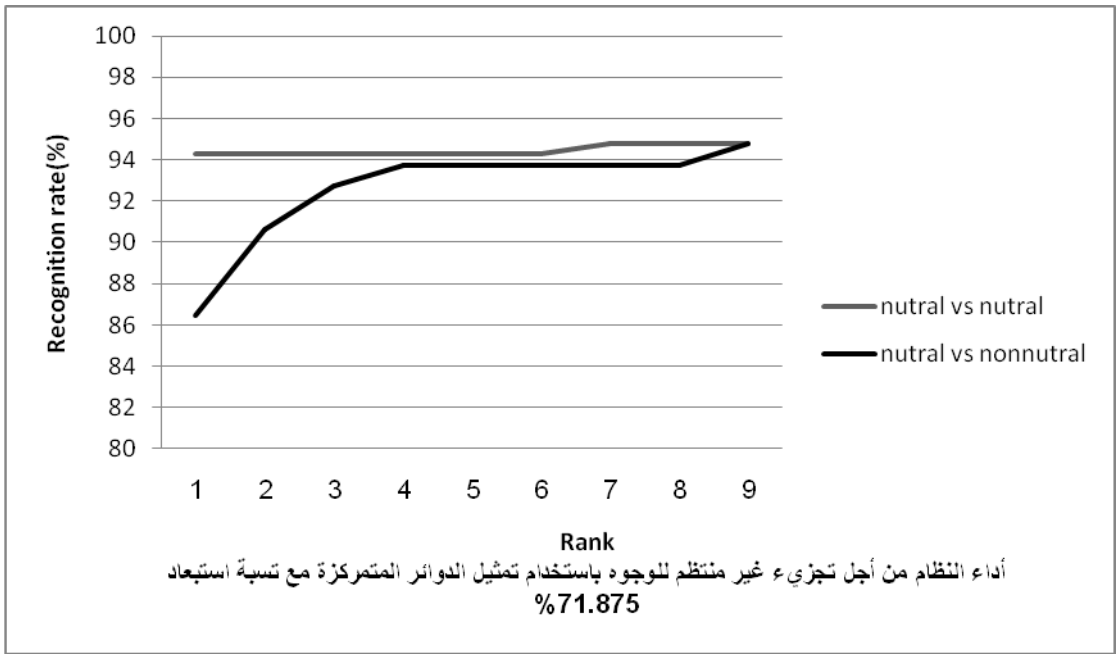
نتائج جيدة مقارنة مع الطرق التي تستخدم PCA عند استخدام PCA للتعرف على الوجوه ثنائية الأبعاد أو ثلاثية الأبعاد وحتى مع استخدام ٥ وجوه للشخص في قاعدة البيانات.

وحققت طريقتنا(بدون تعابير وبدون استبعاد) كذلك نتائج جيدة عند المقارنة مع طرق التعرف المعتمدة فقط على ٣D ICP بدون جمع مع أسلوب تعرف آخر والموضحة في الجداول في الفصل الأول والتي تعتمد على استخدام وجهين أو وجه واحد لكل شخصية في قاعدة البيانات علماً أنه للمقارنة الصحيحية يجب تجربة طريقتنا على نفس قواعد البيانات المستخدمة في الطرق الأخرى التي تستفيد من ICP(حيث تتميز قاعدة البيانات التي استخدمناها بالتنوع في الوضعية pose للوجوه) وكذلك يجب استخدام نفس عدد الوجوه لكل شخص في قاعدة البيانات وغيرها من العوامل والظروف التي يجب أن تتطابق عند المقارنة حيث تؤثر هذه العوامل وغيرها على معدل التعرف.

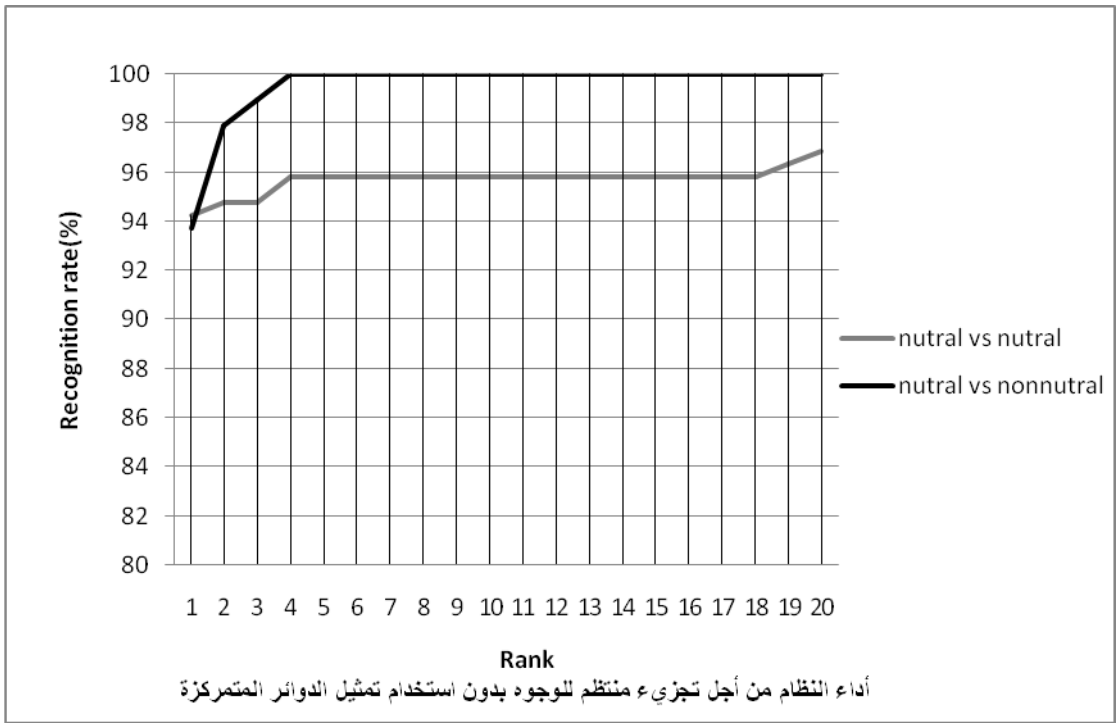
نوضح في الأشكال التالية معدل التعرف Recognition rate من أجل حالات مختلفة ومراتب Ranks مختلفة(منحنيات CMC) حيث يعطي معدل التعرف من المرتبة n معدل التعرف RR عند الأخذ بعين الاعتبار أفضل n تطابق كما وضحنا ذلك في الفصل الأول.



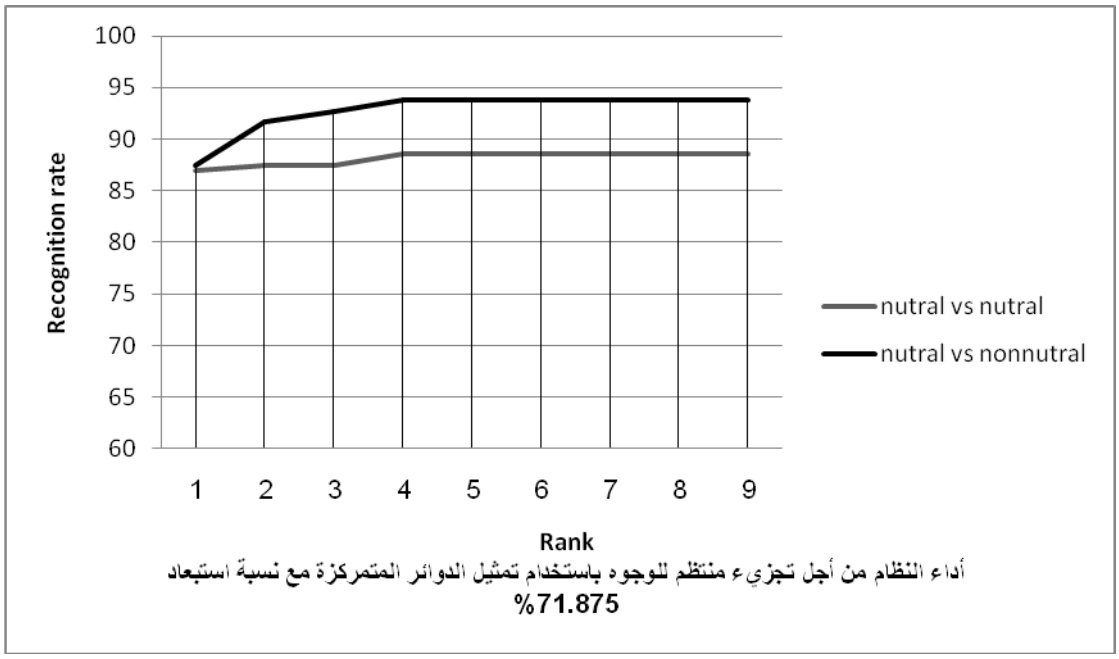
أداء النظام من أجل تجزيء غير منتظم للوجوه بدون باستخدام تمثيل الدوائر المتمركزة



أداء النظام من أجل تجزيء غير منتظم للوجوه باستخدام تمثيل الدوائر المتمركزة مع نسبة استبعاد %71.875



أداء النظام من أجل تجزيء منتظم للوجوه بدون استخدام تمثيل الدوائر المتمركزة



محصلة البحث:

قدم هذا البحث خوارزمية تعرف آلية على الوجوه ثلاثية الأبعاد وضمت المقترحات :

١–تحديد لمنطقة الوجه بواسطة تحديد قمة الأنف

٢–عملية تصحيح آلية لوضعية الوجه

٣–خوارزمية تعرف قوية تجاه التعابير في الوجه بواسطة اقتطاع مناطق الوجه الأقل تأثراً بالتعابير واستخدامها في عملية التعرف

٤–استخدام تمثيل الدوائر المتمركزة للمناطق الأقل حساسية للتعابير وذلك لتسريع الخوارزمية.

العمل المستقبلي:

نلاحظ أن وجود الشعر على الوجه قد يسبب بعض المشاكل لهذه الخوارزمية منها أنه قد يعيق عملية تحديد قمة الأنف لاقتطاع منطقة الوجه وقد يخفض أداء عملية تصحيح الوضعية وكذلك قد يسبب مشاكل أثناء حساب الدوائر المتمركزة لذلك من الممكن مستقبلاً القيام بحذف المناطق المغطاة بالشعر بواسطة خوارزمية skin detection ثم نملأ الفراغات الناتجة. كما يمكن مستقبلاً جمع مزايا التعرف ثنائية الأبعاد مع ثلاثية الأبعاد.

كذلك يمكن مستقبلاً السعي لمعالجة وضعيات الوجه الجانبية profile آلياً والقيام بعملية تعرف آلية ناجحة .

قائمة بالمفردات المستخدمة(عربي-انكليزي):

ازاحة	**translation**
الأشعة الخاصة	**eigenvectors**
الأشعة المفردة	**singular vectors**
الابصار الحاسوبي	**computer vision**
الانحراف المعياري	**standard deviation**
البارامترات الحيوية	**biometrics**
التبعثر	**variance**

التحقق	verification
التحليل إلى القيم المفردة	singular value decomposition
التعرف على النماذج	pattern recognition
التعرف على الوجوه	face recognition
التغاير	covariance
التقنيات المعتمدة على المظهر	appearance based techniques
التقنيات المعتمدة على الملامح المحلية	local features based techniques
التمثيل الدوراني للوجه	spin image
الجار الأقرب	nearest neighbour
متوسط مربع الخطأ الأصغري	least mean squared error
الزائد–الاضافي	rdundancy
الشبكة المثلثية	traingular mesh
الشبكة المربعة	rectangular mesh
الصور الفراغية	range images
الصور ذات التدرج الرمادي	gray–scale images
القيم الخاصة	eigenvalues
القيم المفردة	singular values
الماسح الليزري	laser scanner

الوجه تحت الاختبار	probe face
تجزيء الوجه	face segmentation
تابع الكلفة	cost function
تحديد (أو كشف) منطقة الجلد	skin detection
تحديد (أو كشف) منطقة الوجه	face detection
تحليل المكونات الأساسية	principal component anlaysis
تحويل	transformation
تصحيح الوضعية	pose correction
ثلاثي الأبعاد	٣D(٣ dimensions)
ثنائي الأبعاد	٢D(٢ dimensions)
دوران	rotation
شعاع الازاحة	translation vector
غيمة النقاط	point cloud
قاعدة بيانات الوجوه	gallery faces
مصفوفة الترابط المتعارض	cross-correlation matrix
مصفوفة التغاير	covariance matrix
مصفوفة الدوران	rotation matrix
معدل التعرف	recognition rate

معدل الخطأ المتساوي	equal error rate
معدل الرفض الخاطئ	false rejection rate
معدل القبول الخاطئ	false acceptance rate
منحني خواص عمل-مستقبل	reciever-operating characteristic curve
منحني خواص محصلة التطابق	cumulative match characteristic curve
نسبة الاشارة إلى الضجيج	signal to noise ratio
وضعية أمامية للوجه	frontal pose
وضعية جانبية للوجه	profile pose

قائمة بالمفردات المستخدمة(انكليزي-عربي):

٢dimensions (٢D)	ثنائي الأبعاد
٣ dimensions (٣D)	ثلاثي الأبعاد
appearance based techniques	التقنيات المعتمدة على الملامح
biometrics	البارامترات الحيوية
computer vision	الابصار الحاسوبي
cost function	تابع الكلفة
covariance	التغاير

covariance matrix	مصفوفة التغاير
cross–correlation matrix	مصفوفة الترابط المتعارض
cumulative match characteristic curve	منحني خواص محصلة التطابق
eigenvalues	القيم الخاصة
eigenvectors	الأشعة الخاصة
equal error rate	معدل الخطأ المتساوي
face detection	كشف (أو تحديد)منطقة الوجه
face recognition	التعرف على الوجوه
face segmentation	تجزيء الوجه
false acceptance rate	معدل القبول الخاطئ
false rejection rate	معدل الرفض الخطأ
frontal pose	وضعية أمامية
gallery faces	قاعدة بيانات الوجوه
gray–scale images	الصور ذات التدرج الرمادي
laser scanner	الماسح الليزري
least mean squared error	متوسط مربع الخطأ الأصغري
local features based techniques	التقنيات المعتمدة على الملامح المحلية
nearest neighbour	الجار الأقرب

pattern recognition	التعرف على النماذج
point cloud	غيمة النقاط
pose correction	تصحيح الوضعية
principal component anlaysis	تحليل المكونات الأساسية
probe face	الوجه تحت الاختبار
profile pose	وضعية جانبية
range images	الصور الفراغية
rdundancy	الزائد
reciever–operating characteristic curve	منحني خواص عمل–مستقبل
recognition rate	معدل التعرف
rectangular mesh	الشبكة المربعة
rotation	دوران
rotation matrix	مصفوفة الدوران
signal to noise ratio	نسبة الاشارة إلى الضجيج
singular value decomposition	تحليل القيم المفردة
singular values	القيم المفردة
singular vectors	الأشعة المفردة
skin detection	كشف(أو تحديد) منطقة الجلد

spin image	التمثيل الدوراني للوجه
standard deviation	الانحراف المعياري
traingular mesh	الشبكة المثلثية
transformation	تحويل
translation	ازاحة
translation vector	شعاع الازاحة
variance	التبعثر
verification	التحقق

قائمة بالاختصارات المستخدمة:

٢dimensions (٢D)	ثنائي الأبعاد
٣ dimensions (٣D)	ثلاثي الأبعاد
cumulative match characteristic (CMC) curve	منحني خواص محصلة التطابق
equal error rate(EER)	معدل الخطأ المتساوي
false acceptance rate(FAR)	معدل القبول الخاطئ
false rejection rate(FRR)	معدل الرفض الخطأ
principal component anlaysis(PCA)	تحليل المكونات الأساسية

reciever–operating characteristic (ROC) curve	**منحني خواص عمل–مستقبل**
least mean squared(LMS) error	**متوسط مربع الخطأ الأصغري**
signal to noise ratio(SNR)	**نسبة الاشارة إلى الضجيج**
singular value decomposition(SVD)	**تحليل القيم المفردة**

المراجع :

[١] A Tutorial on Principal Component Analysis
Jonathon Shlens_Systems Neurobiology Laboratory, Salk Insitute for Biological Studies La Jolla, CA ٩٢٠٣٧ and Institute for Nonlinear Science, University of California, San Diego La Jolla, CA ٩٢٠٩٣-٠٤٠٢

[٢] Advances and Challenges in ٣D and ٢D+٣D Human Face Recognition
Shalini Gupta١, Mia. K. Markey٢, Alan C. Bovik١
١ Department of Electrical and Computer Engineering, The University of Texas at Austin, TX ٧٨٧١٢, USA
٢ The University of Texas Department of Biomedical Engineering, Austin, TX ٧٨٧١٢, USA

[٣] GavabDB: a ٣D Face Database
A. B. Moreno, A. Sánchez

Departamento de Informática, Estadística y Telemática, Escuela Superior de Ciencias Experimentales y
Tecnología (Universidad Rey Juan Carlos), c/Tulipán s/n, ٢٨٩٣٣ Móstoles (Madrid), Spain

[٤] K. Arun, T. Huang, and S. Blostein. Least-squares Fitting of Two ٣-D Point Sets. TPAMI, ٩(٥):٦٩٨–٧٠٠, ١٩٨٧.

[٥] R. C. Gonzalez and R. E.Woods. Digital Image Processing. Addison Wesley, ١٩٩٢.

[٦] A. Johnson and M. Hebert. Using Spin Images for Efficient Object Recognition in Cluttered ٣D Scenes. TPAMI, ٢١(٥):٦٧٤–٦٨٦, ١٩٩٩.

[٧] P. Phillips, P. Flynn, T. Scruggs, K. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek. Overview of the Face Recognition Grand Challenge. In CVPR, ٢٠٠٥.

[٨] M. Turk and A. Pentland. Eigenfaces for Recognition. J. Cogn. Neurosci, ٣, ١٩٩١.

[٩] On the ICP Algorithm_Esther Ezra, Micha Sharir, Alon Efrat§

[١٠] K. I. Chang, K. W. Bowyer, and P. J. Flynn. Face recognition using ٢d and ٣d facial data. In Multimodal User Authentication, Workshop in, pages ٢٥–٣٢, ٢٠٠٣.
[١١] A. Colombo, C. Cusano, and R. Schettini. ٣d face detection using curvature analysis. Pattern Recognition, ٣٩(٣):٤٤٤–٤٥٥, ٢٠٠٦.

[١٢] B. Achermann, X. Jiang, and H. Bunke. Face recognition using range images. In Proceedings. International Conference on Virtual Systems and MultiMedia, pages ١٢٩–١٣٦. Int. Soc. Virtual Syst. & MultiMedia (VSMM), Inst. of Comput. Sci. & Appl. Math., Bern Univ., Switzerland, ١٩٩٧.

[١٣] G. Pan, S. Han, Z. Wu, and Y. Wang. ٣d face recognition using mapped depth images. In Computer Vision and Pattern Recognition, ٢٠٠٥ IEEE Computer Society Conference on, volume ٣, pages ١٧٥–١٧٥, ٢٠٠٥.

[١٤] V. Blanz and T. Vetter. Face recognition based on fitting a ٣d morphable model. Pattern Analysis and Machine Intelligence, IEEE Transactions on, ٢٥(٩):١٠٦٣–١٠٧٤،٢٠٠٣.

[١٥] C. Hesher, A. Srivastava, and G. Erlebacher. Principal component analysis of range images for facial recognition. In International Conference on Imaging Science, Systems,and Technology (CISST), in Proceedings of, ٢٠٠٢.

[١٦] K. I. Chang, K. W. Bowyer, P. J. Flynn, and X. Chen. Multi-biometrics using facial appearance, shape and temperature. In Automatic Face and Gesture Recognition,٢٠٠٤. Proceedings. Sixth IEEE International Conference on, pages ٤٣–٤٨, ٢٠٠٤.

[١٧] K. I. Chang, K. W. Bowyer, and P. J. Flynn. Multimodal ٢d and ٣d biometrics for face recognition. In Analysis and Modeling of Faces and Gestures, ٢٠٠٣. AMFG ٢٠٠٣. IEEE International Workshop on, pages ١٨٧–١٩٤, ٢٠٠٣.

[١٨] A. Godil, S. Ressler, and P. Grother. Face recognition using ٣d facial shape and color map information: comparison and combination. In Proceedings of the SPIE – The International Society for Optical Engineering, volume ٥٤٠٤, pages ٣٥١–٣٦١, Nat. Inst. of Stand. & Technol., Gaithersburg, MD, USA, ٢٠٠٤. SPIE-Int. Soc. Opt. Eng.

[١٩] A. Hyvarinen, J. Karhunen, and E. Oja. Independent Component Analysis. John Wiley and Sons, New York, ٢٠٠١.

[٢٠] C. Hesher, A. Srivastava, and G. Erlebacher. A novel technique for face recognition using range imaging. In Signal Processing and Its Applications, ٢٠٠٣. Proceedings.Seventh International Symposium on, volume ٢, pages ٢٠١–٢٠٤ vol.٢, ٢٠٠٣.

[٢١] C. BenAbdelkader and P. A. Griffin. Comparing and combining depth and texture cues for face recognition. Image and Vision Computing, ٢٣(٣):٣٣٩–٣٥٢, ٢٠٠٥.

[٢٢] B. G¨okberk, A. A. Salah, and L. Akarun. Rank-based decision fusion fro ٣d shapebased face recognition. In LNCS, International Conference on Audio- and Videobased Biometric Person Authentication, volume ٣٥٤٦, pages ١٠١٩–١٠٢٨, ٢٠٠٥.

[٢٣] A. Srivastava, X. Liu, and C. Heher. Face recognition using optimal linear components of range images. Accepted for publishing in Image and Vision Computing, ٢٠٠٣.

[٢٤] S. Malassiotis and M. G. Strintzis. Pose and illumination compensation for ٣d face recognition. In Image Processing, ٢٠٠٤. ICIP '٠٤. ٢٠٠٤ International Conference on, volume ١, pages ٩١–٩٤ Vol. ١, ٢٠٠٤.

[٢٥] S. Malassiotis and M. G. Strintzis. Robust face recognition using ٢d and ٣d data:Pose and illumination compensation. Pattern Recognition, ٣٨(١٢):٢٥٣٧–٢٥٤٨, ٢٠٠٥.

[٢٦] F. Tsalakanidou, S. Malassiotis, and M. G. Strintzis. Exploitation of ٣d images for face authentication under pose and illumination variations. In ٣D Data Processing, Visualization and Transmission, ٢٠٠٤. ٣DPVT ٢٠٠٤. Proceedings. ٢nd International Symposium on, pages ٥٠–٥٧, ٢٠٠٤.

[٢٧] F. Tsalakanidou, S. Malassiotis, and M. G. Strintzis. Face localization and authentication using color and depth images. Image Processing, IEEE Transactions on,١٤(٢):١٥٢–١٦٨, ٢٠٠٥.

[٢٨] R. J. Campbell and P. J. Flynn. A survey of free-form object representation and recognition techniques. Computer Vision and Image Understanding, ٨١(٢):١٦٦–٢١٠،٢٠٠١.

[٢٩] B. Achermann and H. Bunke. Classifying range images of human faces with hausdorff distance. In Proceedings ١٥th International Conference on Pattern Recognition, volume vol.٢, pages ٨٠٩–٨١٣. IEEE Comput. Soc, Dept. of Comput. Sci., Bern Univ., Switzerland, ٢٠٠٠.

[٣٠] G. G. Gordon. Face recognition based on depth maps and surface curvature. In SPIE Geometric methods in Computer Vision, volume ١٥٧٠, pages ٢٣٤–٢٤٧, ١٩٩١.

[٣١] M. O. Irfan¨oglu, B. G¨okberk, and L. Akarun. ٣d shape-based face recognition using automatically registered facial surfaces. In Proceedings of the ١٧th International Coneference on Pattern Recognition, volume Vol.٤, pages ١٨٣–١٨٦. IEEE Computer Society, Dept. of Comput. Eng., Bogazici Univ., Turkey, ٢٠٠٤.

[٣٢] K. I. Chang, K. W. Bowyer, and P. J. Flynn. Adaptive rigid multi-region selection for handling expression variation in ٣d face recognition. In Computer Vision and Pattern Recognition, ٢٠٠٥ IEEE Computer Society Conference on, volume ٣, pages ١٥٧–١٥٧, ٢٠٠٥.

[٣٣] X. Lu, D. Colbry, A. K. Jain, and D. Zhang. Matching ٢,٥d scans for face recognition. In D. Zhang and A.K. Jain, editors, Biometric Authentication. First International Conference, ICBA ٢٠٠٤. Proceedings (Lecture Notes in Comput. Sci.Vol.٣٠٧٢), pages ٣٠–٣٦. Croucher Found., Dept. of Comput. Sci. & Eng., Michigan State Univ., East Lasning, MI, USA, ٢٠٠٤.

[٣٤] M. L. Koudelka, M. W. Koch, and T. D. Russ. A prescreener for ٣d face recognition using radial symmetry and the hausdorff fraction. In Computer Vision and Pattern Recognition, ٢٠٠٥ IEEE Computer Society Conference on, volume ٣, pages ١٦٨– ١٦٨, ٢٠٠٥.

[٣٥] Y. Chen and G. Medioni. Object modeling by registration of multiple range images.Image and Vision Computing, ١٠:١٤٥–١٥٥, ١٩٩٢.

[٣٦] P. J. Besl and H. D. McKay. A method for registration of ٣-d shapes. Pattern Analysis and Machine Intelligence, IEEE Transactions on, ١٤(٢):٢٣٩–٢٥٦, ١٩٩٢.

[٣٧] T. D. Russ, M.W. Koch, and C. Q. Little. A ٢d range hausdorff approach for ٣d face recognition. In Computer Vision and Pattern Recognition, ٢٠٠٥ IEEE Computer Society Conference on, volume ٣, pages ١٦٩–١٦٩, ٢٠٠٥.

[٣٨] X. Lu, A. K. Jain, and D. Colbry. Matching ٢,٥d face scans to ٣d models. Pattern Analysis and Machine Intelligence, IEEE Transactions on, ٢٨(١):٣١–٤٣, ٢٠٠٦.

[٣٩] E. P. Kukula, S. J. Elliott, R. Waupotitsch, and B. Pesenti. Effects of illumination changes on the performance of geometrix facevision/spl reg/ ٣d frs. In Security Technology, ٢٠٠٤. ٣٨th Annual ٢٠٠٤ International Carnahan Conference on, pages ٣٣١–٣٣٧, ٢٠٠٤.

[٤٠] M. O. Irfan¨oglu, B. G¨okberk, and L. Akarun. ٣d shape-based face recognition using automatically registered facial surfaces. In Proceedings of the ١٧th International Coneference on Pattern Recognition, volume Vol.٤, pages ١٨٣–١٨٦. IEEE Computer Society, Dept. of Comput. Eng., Bogazici Univ., Turkey, ٢٠٠٤.

[٤١] G. Medioni and R. Waupotitsch. Face modeling and recognition in ٣-d. In Analysis and Modeling of Faces and Gestures, ٢٠٠٣. AMFG ٢٠٠٣. IEEE International Workshop on, pages ٢٣٢–٢٣٣, ٢٠٠٣.

[٤٢] T. Maurer, D. Guigonis, I. Maslov, B. Pesenti, A. Tsaregorodtsev, D. West, and G. Medioni. Performance of geometrix activeidTM ٣d face recognition engine on the frgc data. In Computer Vision and Pattern Recognition, ٢٠٠٥ IEEE Computer Society Conference on, volume ٣, pages ١٥٤–١٥٤, ٢٠٠٥.

[٤٣] T. Papatheodorou and D. Rueckert. Evaluation of automatic ٤d face recognition using surface and texture registration. In Automatic Face and

Gesture Recognition, ٢٠٠٤. Proceedings. Sixth IEEE International Conference on, pages ٣٢١–٣٢٦, ٢٠٠٤.

[٤٤] T. Nagamine, T. Uemura, and I. Masuda. ٣d facial image analysis for human identification. In Pattern Recognition, ١٩٩٢ . Vol.١. Conference A: Computer Vision and Applications, Proceedings., ١١th IAPR International Conference on, pages ٣٢٤–٣٢٧, ١٩٩٢.

[٤٥] C. Beumier and M. Acheroy. Automatic ٣d face authentication. Image and Vision Computing, ١٨(٤):٣١٥–٣٢١, ٢٠٠٠.

[٤٦] C. Beumier and M. Acheroy. Face verification from ٣d and grey level clues. Pattern Recognition Letters, ٢٢(١٢):١٣٢١–١٣٢٩, ٢٠٠١.

[٤٧] L. Zhang, A. Razdan, G. Farin, J. Femiani, M. Bae, and C. Lockwood. ٣d face authentication and recognition based on bilateral symmetry analysis. Visual Computer, ٢٢(١):٤٣–٥٥, ٢٠٠٦.

[٤٨] I. A. Kakadiaris, G. Passalis, T. Theoharis, G. Toderici, I. Konstantinidis, and N. Murtuza. Multimodal face recognition: combination of geometry with physiological information. In Computer Vision and Pattern Recognition, ٢٠٠٥. CVPR ٢٠٠٥. IEEE Computer Society Conference on, volume ٢, pages ١٠٢٢–١٠٢٩ vol. ٢, ٢٠٠٥.

[٤٩] M. L. Koudelka, M. W. Koch, and T. D. Russ. A prescreener for ٣d face recognition using radial symmetry and the hausdorff fraction. In Computer Vision and Pattern Recognition, ٢٠٠٥ IEEE Computer Society Conference on, volume ٣, pages ١٦٨– ١٦٨, ٢٠٠٥.

[٥٠] C.-S Chua, F. Han, and Y.-K. Ho. ٣d human face recognition using point signature. In Automatic Face and Gesture Recognition, ٢٠٠٠. Proceedings. Fourth IEEE International Conference on, pages ٢٣٣–٢٣٨, ٢٠٠٠.

[٥١] B. G¨okberk, A. A. Salah, and L. Akarun. Rank-based decision fusion fro ٣d shapebased face recognition. In LNCS, International Conference on Audio- and Videobased Biometric Person Authentication, volume ٣٥٤٦, pages ١٠١٩–١٠٢٨, ٢٠٠٥.

[٥٢] G. Pan, Y.Wu, and Z.Wu. Investigating profile extracted from range data for ٣d face recognition. In Systems, Man and Cybernetics, ٢٠٠٣. IEEE International Conference on, volume ٢, pages ١٣٩٦–١٣٩٩ vol.٢, ٢٠٠٣.