تجزئة الأغراض في الصور الرقمية وتصنيفها[1]

م. المهندس محمد أيمن العقاد[2] الدكتور المهندس حسن أبو النور[3]

الملخص

نعرض في هذه الورقة تحديداً آلياً لنقاط العلام المسيطرة dominant landmarks في الأجسام وتقديراً لتقوسات سطوح الأجسام objects surfaces curvature بهدف الوصول إلى موصفات شكل ثابتة invariant shape descriptions. إنَّ الموصفات المستنبطة ضرورية من أجل بناء نماذج الأجسام ثلاثية الأبعاد وتحليلها وعمليات التعرف. تُستخدم المعرفة المسبقة a priori knowledge ببنية الوجه البشري. استخدمنا طريقة جديدة في تمثيل الأجسام، دعوناها صيغة عرض النبضة المتغير(varying pulse width form) VPWF، والتي ليست فقط مناسبة من أجل التعرف بل أيضاً في التنفيذ العتادي، والاتصالات، وتقنيات الضغط، والتزامن، والرسوم المتحركة.

كلمات مفتاحية: نقاط العلام المسيطرة landmarks dominant، تقوسات السطوح Surface curvature، المعرفة المسبقة a priori knowledge ببنية الوجه البشري، بناء النماذج ثلاثية الأبعاد 3D model construction، تعرّف الأشكال Pattern recognition.

[1] أعد هذا البحث في سياق رسالة الدكتوراه للطالب المهندس محمد أيمن العقاد بإشراف الأستاذ الدكتور المهندس حسن أبو النور.

[2] قسم هندسة الحواسيب والأتمتة– كلية الهندسة الميكانيكية والكهربائية– جامعة دمشق.

[3] قسم هندسة الالكترونيات والاتصالات– كلية الهندسة الميكانيكية والكهربائية– جامعة دمشق.

1. مقدمة:

بناءً على خريطة التوجه، المستنبطة في [1]، والتي تؤمن نقلاً للصورة رمادية التدرج إلى حقل شعاع الناظم normal، يمكننا استخراج نقاط العلام المسيطرة في الأجسام وتقوسات السطوح.

نقدم في هذا البحث تقنية جديدة دعوناها صيغة عرض النبضة المتغير VPWF لتمثيل الأغراض Variable Pulse Width Format، تمتلك هذه التقنية مزايا كثيرة بالنسبة للطرائق الموجودة [15]،[16]،[18] فيما يتعلق بالتنفيذ، وضغط المعطيات، والتعامل مع الرسوم المتحركة.

سنحدد نقاط علام مسيطرة، والتي هي ثابتة بالنسبة لتغيرات الوضعية وشروط الإضاءة، والتي تبقى ثابتة بالنسبة لكل أنواع التحويلات. سيتم التحديد بكشف تلك العبورات المصاحبة لتغيرات في محتوى الصورة أو التباين، وعلى أساس البنية المعروفة للوجه البشري. من ثم تثبت نقاط العلام هذه ضمن النموذج ثلاثي الأبعاد، وتستخدم لتحديد النقاط الأساسية لسطوح الجسم.

من ثم نستخدم تقدير التقوس للسطوح، لأن تقوسات السطوح هي المفتاح في حساب موصّفات الأشكال وتصنيف أنواع مختلفة من السطوح. كما أن موصفات التقوس للوجه البشري غنية بالمعلومات، وفريدة مثل بصمات الأصابع. تقدم تغيرات التقوس توصيفاً مستقراً نسبياً، لأنها تبقى قابلة للكشف ضمن مجال واسع من التحويلات الإحداثية. (نستخدم الزوايا بين أشعة متعامدات رقع سطح متجاورة في قياس التقوس المنفصل).

تملك الأشكال المتشابهة في المفهوم العام موصفات متشابهة، حتى لو اختلفت في التفاصيل ضيقة المجال. كما تبنى الأشكال المعقدة من ترتيبات معينة لمكونات أساسية.

عند الحديث عن الأجسام ثلاثية الأبعاد لابد من التطرق إلى المحددات التفاضلية التي تصف هذه الأجسام (كزوايا الأعمدة على السطوح و تقوسات السطوح) حيث توجد طرائق لحساب هذه المحددات وتقديرها في جوار محدد.

وتقسم الأجسام ثلاثية الأبعاد إلى:

1. أجسام ثلاثية أبعاد بسيطة [4]وهي (إما رقع سطوح مستوية Planar، أو رقع سطوح كروية Spherical، أو رقع سطوح اسطوانية Cylindrical) وتستخدم في بناء نماذجها الهندسية (إما كثيرات الأضلاع Polygons [5]، أو ثنائيات الدرجة المتعددة Super-quadrics [6]، أو التقسيم المثلثي Triangulation).

2. أجسام ثلاثية أبعاد معقدة وتستخدم في دراستها الطرائق الآتية (إما طرائق تحليل الأشكال Shape decomposition [7,8]، أو البنى الخطية Linear object classes [9]، أو تمثيل الأجسام المعقدة Complex Objects Representation [10]) وتستخدم في بناء نماذجها الهندسية (إما النماذج غرضية التوجه Object-Oriented models [12]، أو النماذج الطبوغرافية Topological models [11]، أو التحليل الفراغي [13,14] Spatial analysis).

2. تحديد نقاط العلام المسيطرة:

تعريف(1): نقاط العلام المسيطرة في جسم ما هي نقاط ذات تقوس كبير أو نقاط وصل.

وتحدد أي ثلاث نقاط تقع في مستوي ولا تقع على مستقيم واحد شعاعين، هما الأساس في تحديد مواقع نقاط أخرى تقع في المستوي نفسه.

تحسب زوايا توجه الناظم على عقد nodes النموذج ثلاثي الأبعاد بالنسبة للشعاع Sp الناظم على المستوي المرجعي (مستوى الصورة)، وتختار الخطوة المنتظمة في حساب النموذج ثلاثي الأبعاد بشكل يتم فيه الحفاظ على تفاصيل الجسم. تحدد العقدة في نموذجنا ثلاثي الأبعاد نقطة، أو نقطة نهاية، أو زاوية في مضلع حيث تلتقي

حافتان. يكون التشكل الناتج ثابتاً بالنسبة لتغيرات الوضعية، ومن ثَـمَّ ثابتـاً بالنـسبة لتحولات منظورية طفيفة. إن نقاط العلام الخاصة في الوجه يمكن أن تكون غمازة، أو وحمة، أو شامة، أو بقعة جلدية، أو خالاً، أو صباغاً جلدياً، أو ندبة. في حـين نقـاط العلام الأساسية هي زوايا العيون، رأس الأنف، مركز الفم، الذقن، الخ.

تظهر مجموعات من التقعرات شبه المتناظرة عند الزوايا الداخلية للعيـون أو جـسر الأنف، والزوايا الخارجية للعيون، وقاعدة الأنف وغالباً ما تظهر أيضاً عند المنخرين، وعند الزوايا الخارجية للفم، وعند الغمازات. من ضمن هذه النقاط وجدنا أن زوايا الفم والغمازات غالباً ذات درجة تناظر أقل من البقية. إن التقوسين الأساسـيين يظهـران بوضوح كمناطق محدبة شبه متناظرة تحدد موقع المقلتين والجفنين.

إن أهم متطلبات التعرف الآلي هي تحديد نقاط العلام الثابتة بالنـسبة لتعـابير وجـه مختلفة والتي تتم من خلال إيجاد القيم المحلية الصغرى والعظمى عبر تحديد النقـاط ذات التغير في الإشارة من موجب إلى سالب ومن سالب إلى موجب للزاوية α والتي تحقق شرط عتبة محدداً علماً بأن دخل البرنامج هو صورة ثنائية البعد. وقد وجدنا أن الوجوه تميل إلى عدم امتلاك مناطق مسطحة كبيرة أو مناطق كروية، بل بدلاً من ذلك تمتلك نقاطاً مفروزة محدبة، والتي تُعلّم صفات مميزة في خطوط التقوس الأساسي.

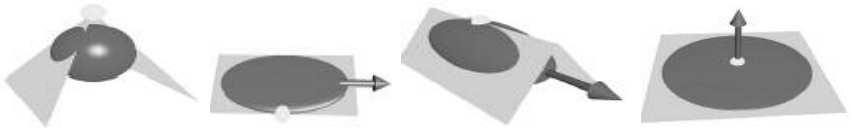
بمساعدة جهة الناظم، يمكننا استنباط تابع تحديد للحافة. ومن أجل ذلك، نسقط كل عقدة V مع جوارها على نظام إحداثيات xy متعامد مع جهة الناظم، حيث V هي المبدأ. ثم نفرز الجوار حسب زاويتها حول V. لندع β أكبر فرجة زاوية والتي لا تحـوي أي نقطة جوار. فكلما كانت β أكبر كان احتمال كون V نقطة حافة أكبر. وكنتيجة، نميز بين ما يأتي:

عقد حافة: عند نقاط الحافة.

عقد تجعيد: تنتهي عند وصلات، أو نقاط نهاية، أو تنغلق لتشكل حلقة.

عقد نهاية: عند نهايات وحيدة النقطة.

الرقع حول عقد كهذه مبينة في الشكل 1.



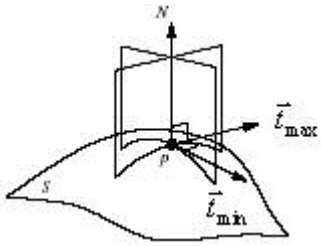
(د) (ج) (ب) (آ)

الشكل 1. شكل الرقع حول العقد: آ) عقدة سطح، ب) عقدة تجعد، ج) عقدة حافة، د) عقدة زاوية.

3. تقدير تقوسات السطوح:

تعريف(2): التقوس هو تغير في التوجه.

تقدير التقوس يستغل البنية التفاضلية للصور [3]. فعند كل نقطة P على السطح يتشكل منحني من تقاطع السطح والمستوي الناظم في اتجاه محدد $\vec{t}_i$. تقوس هذا المنحني المستوي هو التقوس العمودي k_n عند P في الاتجاه المعطى $\vec{t}_i$. إن التقوسات العمودية العظمى والصغرى عند نقطة ما تحدد التقوسين الأساسيين k_{max} وk_{min}، والاتجاهين $\vec{t}_{\max}$ و $\vec{t}_{\min}$ المصاحبين لكل من التقوسين الأساسيين هما الاتجاهان الأساسيان للسطح عند P وتبعاً لنظرية أولر هما متعامدان ويحددان مع شعاع الناظم نظاماً متعامداً كما يبين الشكل 2. يتحدد التقوس الوسطي بالعلاقة $H=(k_{min}+k_{max})/2$. ويتحدد تقوس غاوص بالعلاقة $K=k_{min}.k_{max}$.



الشكل 2. التقوسان الأساسيان k_{max} وk_{min} بالاتجاهين الأساسيين $\vec{t}_{\max}$ و $\vec{t}_{\min}$.

يصف دليل الشكل (*S*) *Shape index*، وهو معيار عام لتقوس السطح وتقعره، الشكل المحلي عند نقطة بشكل مستقل عن قياس السطح [4]، ويتحدد بالعلاقة:

$$S = -\frac{2}{\pi} \arctan \frac{k_{\max} + k_{\min}}{k_{\max} - k_{\min}} \quad , \quad -1 \le S \le 1.$$

تمتلك نقطة سطح محدب ذي تقوسات أساسية متماثلة $S=1$، و نقطة سطح مقعر ذي تقوسات أساسية متماثلة القيمة $S=-1$، ونقطة سطح سرجي ذي تقوسات أساسية متماثلة بالقيمة متعاكسة بالإشارة $S=0$، أما لنقطة سطح على شكل هضبة فلها $S=0.5$، بينما لنقطة سطح على شكل وادٍ تكون لها القيمة $S=-0.5$، ويبين الشكل 3 أمثلة لأشكال هذه السطوح.

الشكل 3. أمثلة لأشكال السطوح على التوالي من اليسار إلى اليمين (سطح مقعر، وسطح محدب، وسطح على شكل هضبة، وسطح على شكل وادٍ، وسطح سرجي)

إن مقدار الانحناء هو معيار آخر يصف قياس السطح بشكل مستقل عن شكله $R = \sqrt{k_{\min}^2 + k_{\max}^2}/2$، حيث النقاط المسطحة ذات $R=0$ في حين المناطق المنحنية بشدة ذات R كبيرة القيمة.

رقع السطح المسطحة: عندما تكون قيمة كل من التقوسين الأساسيين صفراً لنقاط السطح يدعى سطحاً مسطحاً.

رقع السطح شبيه الكرة: هي رقع حيث تمتلك النقاط فيها قيمة التقوسات الأساسية نفسها. يمكن فصلها مبدئياً بتحديد النقاط التي تحقق $|k_{max}-k_{min}|<e$، حيث e هي قيمة عتبة محددة.

رقع السطح الاسطوانية: هي رقع حيث النقاط فيها على شكل قطع مكافئ (أي تقوس غاوص يساوي الصفر) وبتقوس وسطي غير صفري.

سطح شبيه الكرة صغير يحوي المعلومات نفسها كسطح شبيه الكرة كبير حيث الفارق هو القياس وليس المعلومات، من ثَمَّ قيمة تقوس أكبر هي معيار مستوى وليس معيار معلومات. إن دليلاً غير صفري لحقل الاتجاه حول نقطة هو مؤشر أكثر وثوقيةً لنقطة

شبيهة الكرة لأنه يأخذ بالحسبان الجوار حول النقطة بدلاً من قيم التقوس عند نقطة وحيدة فقط. فضلاً عمَّا سبق نجد أن التدرج، وهو عملية تفاضلية، يُمكّننا من اشتقاق الشكل المحلي ومعرفة توجه السطح في المشهد. ومن ثَمَّ فإننا سنستخدم تدرج زوايا توجه النموذج ثلاثي الأبعاد α حيث تدرج الزاوية α في اتجاه أساسي متناسب مع التقوس الأساسي في هذا الاتجاه:

$$\alpha_{max} = const\ .\ k_{max}\ ,\ \alpha_{min} = const\ .\ k_{min} \qquad (1)$$

1.3. العناصر الأولية للأشكال:

إن الشكل هو المحل الهندسي لمجموعة نقاط تقع على سطح الجسم الفيزيائي، أو على حافاته أو لمنطقة ذات شدة إضاءة معينة في صورة.

تعريف(3): العنصر الأولي لشكل هو مجموعة من عقدة أو أكثر، وهو نفسه ما ندعوه معالم الجسم.

تعريف(4): الموصف هو كينونة، تمثل عنصراً أولياً للشكل (ميزة بارزة في الجسم).

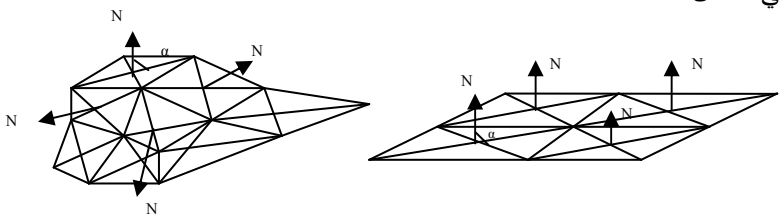
يقيس معامل الشكل كيف يتغير ناظم السطح N لدى التحرك باتجاهات مختلفة في الفضاء المماس للسطح، كما أن معلومات الشكل تمثل قيمة انتروبي[1] التقوس curvature entropy. يوصف الشكل بتركيبة من العناصر الأولية حيث كل عنصر يصف جزءاً محدداً من الشكل الإجمالي. تمكن المعالم المكتشفة من تجزيء مسألة إعادة بناء السطح إلى مسائل فرعية أبسط على رقع سطح ناعم. لكل شكل هناك وصف فريد، وإذا كانت الأشكال متشابهة على نطاق واسع، فلها موصفات متشابهة، حتى لو اختلفت في نطاق ضيق (تفاصيل). تُصنّف مناطق السطوح إلى مسطحة، اسطوانية، بيضاوية، أو زائدية المقطع. نقارن عادةً الأشكال لنحكم هل هي متشابهة بشكل كافٍ وليست متطابقة، مع بعض التسامح تبعاً للضجيج والتغطية. تقرب

[1] الانتروبي هو معيار التأكد من غرض بعد أخذ خصائصه بالحسبان وهو تابع احتمالي للحالات التي يكون عليها الغرض، كلما زاد الانتروبي زاد عدم التأكد أي زادت الحالات وزاد احتمال هذه الحالات.

السطوح بسلسلة من n رقعة، حيث تتألف الأجسام المعقدة من تشكيلات من أجزاء أساسية. فيملك السطح بضعة معالم عند دقة تمييز منخفضة في حين تظهر معالم عديدة عند دقة تمييز عالية. ويفضل عادةً تجاهل المعالم العالية التمييزية وتحليل الخصائص العامة للشكل فقط.

توجد تقنيتان لحساب خصائص السطح: [8] تقنية حجمية الأساس معتمدة على تابع كثافة تقريب، حيث التحليل فيها يتم على سطح منخفض التمييزية نُعِّمَتْ معالمه عالية التمييزية بواسطة طي الكثافة مع تابع غاوص ثلاثي الأبعاد وذي تعارض مناسب، وتقنية سطحية الأساس، والتي توسط فيها الخصائص ضمن رقعة السطح بدلاً من التوسيط ضمن الحجم.

يوصف الشكل المحلي عند كل نقطة من السطح بواسطة عدة كميات سلمية وشعاعية، والمقيم الخاص بنا يعمل على حقل التوجه (شعاع الناظم) للسطح. ونظراً لأنَّ لدينا مسحاً منتظماً على كامل السطح، فإنَّ التقوس متناسب مباشرةً مع زوايا التوجه. سنحلل كامل سطح الجسم بناءً على عناصر رقعة حيث تحتوي كل رقعة أربع عقد كما في الشكل 4.



شكل 4. عناصر رقعة سطح جسم.

نحسب $\Delta\alpha1=\alpha1-\alpha2$، $\Delta\alpha2=\alpha1-\alpha3$، $\Delta\alpha3=\alpha3-\alpha4$، و $\Delta\alpha4=\alpha2-\alpha4$، عندها تكون الرقعة مسطحة إذا كان $\Delta\alpha1=\Delta\alpha2=\Delta\alpha3=\Delta\alpha4=0$، والرقعة شبه الكرة إذا كان $\Delta\alpha1=\Delta\alpha2=\Delta\alpha3=\Delta\alpha4\neq0$، والرقعة اسطوانية في أحد الاتجاهات الأساسية إذا كان $\Delta\alpha1=\Delta\alpha2\neq0$ و $\Delta\alpha3=\Delta\alpha4=0$، واسطوانية في الاتجاه

الأساسي الآخر إذا كان $\Delta\alpha 1 = \Delta\alpha 2 = 0$ و $\Delta\alpha 3 = \Delta\alpha 4 \neq 0$. في المناطق حيث لا توجد تغيرات كبيرة في زاوية التوجه، نحفظ في قاعدة المعطيات فقط عقد الحافات وإحداثياتهم من أجل تخفيض حيز التخزين اللازم. نحسب معلومات الأشكال باستخدام الانتروبي:

$$E(\alpha) = -\sum_{i} p_i \log p_i$$

حيث p_i تابع كثافة الاحتمال المنفصل للمتحول العشوائي α.

سنحدد لكل نوع وجهاً (مثلاً بيضوي) ونوع معالم (مثلاً شفاه ممتلئة) موصف الشكل الخاص بها. تميز قيم الموصفات المختلفة بين ثلاث مجموعات وجوه رئيسية مختلفة تبعاً للعمر، والجنس، والعرق، والتي تتنزع المميزات التفردية. إن قيم الزوايا الكبيرة تتعلق بالوصف العام، في حين قيم الزوايا الصغيرة تتعلق بالتفاصيل الأدق. من ثَمَّ يتم القرار عن العرق، والجنس، والعمر باستخدام قيم زوايا كبيرة، في حين تُحَدَّدُ الهوية الفردية يتم باستخدام قيم الزوايا الصغيرة. سيتم ذلك من خلال مسح النموذج ثلاثي الأبعاد في الاتجاهين الأساسين، وهما اتجاه المحور x و اتجاه المحور y.

2.3. تنظيم العناصر الأولية للأشكال:

تعريف(5): تنظيم العناصر الأولية للأشكال هو عبارة عن تمثيل الشكل كشبكة بيانية (أو شبكة فوق البيانية **hyper-graph**) حيث ترمز العقد إلى عناصر أولية أو مجموعات من عناصر أولية، والأقواس ترمز إلى العلاقة بينهم.

إن القرار عن تجميع سمات محددة يتم نموذجياً بقياس أشكال مختلفة من الهندسة الفراغية النسبية المتعلقة بهم وبمقارنة هذه القياسات مع قيم عتبات محددة. إذا كانت قيمة زاوية ميل الناظم لعقدتين قريبة إلى حد كافٍ، ولا توجد أية زاوية بارزة بينهما يتخذ قرار بتجميعهما، مولدين سمة جديدة ترمز إلى عقدتين تقعان على المستوي نفسه. ويجب أن لا تتجاوز أية مجموعة حدود الجسم. نتيجة المطابقة من المستوى الأعلى تؤمن نقطة بداية متقدمة في البحث عن مطابقة أكثر اكتمالاً. إن هذا يتطلب

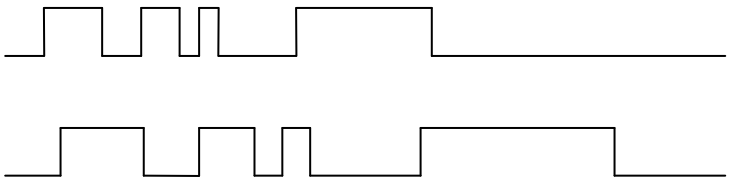
ثلاث سمات نقطية إذا كان الجسم صلباً، وأقل إذا كانت للسمات النقطية توجيهات مصاحبة (مثلاً كما للوصلات)، وأكثر إذا كان الجسم قيد التعرف نوعاً ما مرناً. إن البنية العلائقية هي مجموعة من شبكتين بيانيتين، شبكتان ثنائيتا العقدة تمثلان عنصرين هندسيين أوليين وعلاقات التوضع بينهما (اتصالية العقد). تستخلص عملية الفهرسة العلائقية السمات، تحولها إلى شبكتين بيانيتين، وتستخدمهما للفهرسة ضمن قاعدة معطيات النماذج ويستخدم التصويت من أجل أصناف المشهد لتلك الأجسام ذات البنية المتشابهة.

يتحدد الشكل داخل الجسم بإحدى طريقتين، محددة الحافات (الداخل ضمن مجموعة متصلة من النقاط ذات قيمة محددة)، أو محددة الداخل (الداخل هو أكبر مجموعة من النقاط المتصلة الحاوية النقطة المولدة والتي لها قيمة هذه النقطة نفسها). أي منطقة متصلة يمكن أن تكون إما شبكة رباعية الاتصال 4-connectivity (أي نقطتين تتصلان إذا كان بينهما ممر بالتحرك أعلى أو أسفل أو يساراً أو يميناً)، أو شبكة ثمانية الاتصال 8-connectivity (أي نقطتين تتصلان إذا كان بينهما ممر بالتحرك أعلى أو أسفل، يساراً أو يميناً أو أعلى– يساراً أو أعلى– يميناً أو أسفل– يساراً أو أسفل– يميناً).

3.3. التمثيل بواسطة عرض النبضة المتغير:

حفظنا في قاعدة المعطيات لكل صورة بنية معطيات للنموذج ثلاثي الأبعاد، و التي تتضمن إحداثيات العقد وزاوية ميلها. لاحقاً عندما نأخذ صورة للتعرف، نحول الصورة إلى نموذجها ثلاثي الأبعاد ومخطط زواياها. ومن ثم نقارن هذا النموذج مع النماذج المثبتة في قاعدة المعطيات، فإذا لم نجد تطابقاً نمرر النموذج عبر إجراء دوران بزاوية محددة حول المحاور x، y، z على التوالي تبعاً لمنطق محدد. نجري المقارنة بشكل متكرر باستخدام تقنية أسميناها تقنية حلقة مطابقة الجسم (OML) Object Matching Loop ، والتي تحسب $\Delta\theta$، وعندما $\Delta\theta = \delta$ (حيث δ قيمة ثابتة

تنتهي إلى الصفر) هذا يعني أن المطابقة قد حصلت وبتقريب معين. سنمثل الجسم بشكل حيث قيم زوايا التوجه للنموذج ثلاثي الأبعاد، المحسوبة باستخدام طريقة حساب توجه الجسم [1]، تحول إلى قيم عرض نبضة كما في الشكل 5، تقنية دعوناها صيغة عرض النبضة المتغير (VRWF) للأجسام.



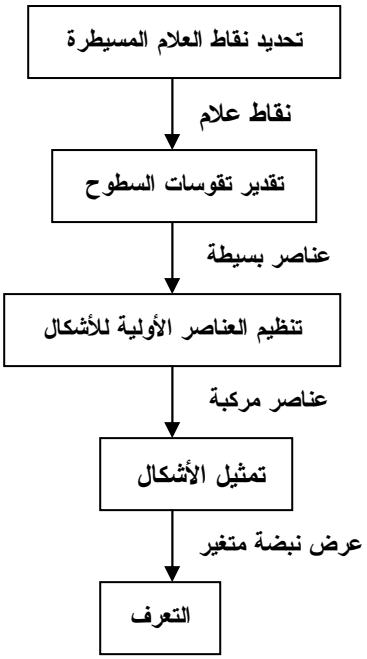
الشكل 5. التمثيل بواسطة عرض النبضة المتغير من أجل شريطين لصورة وجه كمثال.

إن هذا التمثيل مناسب جداً لاستخدامه في تطبيقات مختلفة مثل تقنيات الضغط، أو في الرسوم المتحركة وخصوصاً في حل مسألة تزامن الصورة مع الكلام، أو في تأمين التزامن في الاتصالات الحديثة [2].

4.3. النظام البرمجي:

يتألف النظام البرمجي من الإجرائيات المبينة على المخطط الانسيابي في الشكل 6. كما بُنِيَتْ قاعدة المعطيات الأولية لتتألف من N نموذج وجه مخزن، حيث $N=252$ محسوبة أخذاً بالحسبان *2* مجموعة جنس، و *3* مجموعات عمر، و وجهاً حيادياً، و *6* تعابير عاطفية، و*3* مجموعات عرق، و*2* اتجاهاً للوجه.

ندخل إلى البرنامج صورة ثنائية البعد للوجه ونحصل في الخرج على نسبة تطابق الوجه مع الوجوه المخزنة في قاعدة المعطيات، حيث تظهر النتيجة خلال *0.8* من الثانية تقريباً بحال أُجْرِيَتْ عمليات المعالجة المسبقة preprocessing، أما مقدار الزمن الكلي مضافاً إليه زمن المعالجة المسبقة فهو *4.2* دقيقة تقريباً.



الشكل6. المخطط الانسيابي للبرنامج.

4. التعرف:

يتم التعرف على أي جسم بواسطة سماته (العناصر الأولية للشكل)، ونموذج الجسم يحدد هذه الخصائص. عند التعرف على جزء فرعي، نستخدم أولاً السمات الخشنة من ثم نتابع إلى الأنعم منها. بتحليل البحث بهذه الطريقة، يستبدل بحث كبير ببحوث كثيرة أصغر مما يقود إلى حسابات أقل. تبنى الفهرسة على التجزيء الهندسي، والمطابقة على الرصف. يتم تجزيء الجسم إلى مقاطع مختلفة، والتي تصبح بدورها العناصر الأولية المفهرسة والتي على ضوئها يتم التعرف على جسم محدد. بدلاً من تخزين

كامل المعطيات الهندسية لكل جسم، تخزن فقط المؤشرات المفهرسة لأجزاء الجسم المختلفة.

1.4. الوجه الوسطي:

الوجه الوسطي $r_0(\theta,t)$ هو متوسط الوجوه الموجودة في قاعدة المعطيات الأولية التي احتوت 108 وجه.

عندها أي وجه يكون $r(\theta,t)$ والذي هو عبارة عن تبديل حول الوجه الوسطي حيث $\rho(\theta,t)$ هي التغيرات التي تلتقط هوية الشخص (حيث $\rho/r_0 < 1$).

$$r(\theta,t)= r_0(\theta,t)+ \rho(\theta,t) \quad : \quad r_0=1/N\sum_{k=1}^{N} r^k(\theta,t), \quad \rho=\sum_i \alpha_i \ \psi_i(\theta,t) \tag{2}$$

حيث حسبت قيمة $N=252$ وفقاً للاعتبارات المذكورة في الفقرة (4.3.).

يطبق عامل الزمن t عند وجود إما تغيرات آنية (رسوم متحركة) أو تغيرات طويلة الأمد (نمو وهرم)، حيث θ متناسبة مع t.

2.4. بنية المعطيات:

من الضروري استخدام بنى معطيات فراغية لإيجاد الأجسام ثلاثية الأبعاد ضمن منطقة محددة بأقصى سرعة. البنى الآتية مناسبة للاستخدام من أجل هذا الغرض.

شجرة الصناديق المتداخلة: تتألف من صناديق ثلاثية الأبعاد متداخلة والتي تحوي إما أجساماً هندسيةً أو صناديقَ المستوى اللاحق من الشجرة. تجمع الأجسام تبعاً لنقطة مركز الصندوق المحيط بها. وبهذه الطريقة من المستحيل أن تصبح الشجرة غير متوازنة.

شجرة الصناديق غير المتداخلة: هي مصفوفة حجمية بعقد ثنائية على أساس مبدأ التحليل العودي للفراغ. الصناديق هنا غير متداخلة. كل عقدة تشير إلى منطقة هندسية واحدة بارزة، ولا يمكن لأي عقدة أخرى أن تغطي المنطقة نفسها. إن بنية المعطيات

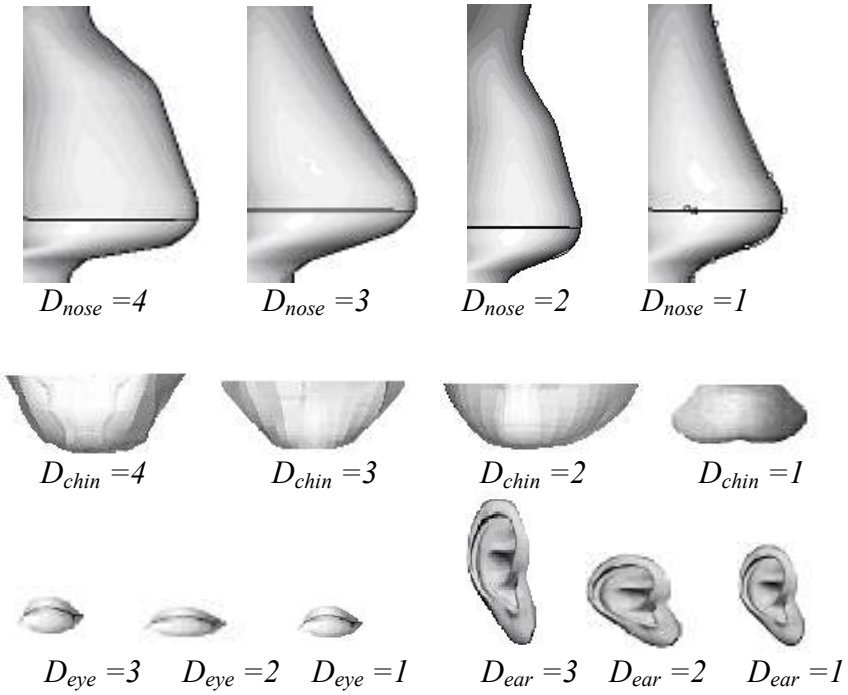
الخاصة بنا تسمح بتخزين ليس الحجم الكامل للجسم، ولكن فقط العقد (نقاط السطح) الضرورية، والتي دعوناها شجرة نقاط السطح.

5. النتائج التجريبية:

الوجه البشري هو الجزء الرئيسي من الرأس المستخدم للتعرف ومبني بطريقة هرمية كما هو مبين في الشكل 7. الأجزاء الرئيسية معلمة بقطع ناقص ومناطق التجاعيد مبينة بخطوط، وكل جزء مع أجزائه الفرعية لها نظام إحداثياتها الهرمي الخاص. هذه البنية تجعل حساب معاملات الوجه أسهل. إذْ على سبيل المثال يتم فقط تغيير نظام إحداثيات الفم عندما يتحرك الفم.

الشكل 7. البنية الهرمية للوجه البشري.

لدى تحليلنا للوجوه وتجزيئها إلى مكوناتها الأساسية استطعنا تصنيفها بناءً على تصنيف هذه المكونات ويبين الشكل 8 مثالاً لمجموعة مختلفة من نماذج الأنف ورأس الذقن والأذن وحجرة العين التي حصلنا عليها والتي أسندنا إلى كلّ نموذج منها موصفاً تصنيفياً خاصاً $D_{nose}=\{1,2,3,4\}$، $D_{chin}=\{1,2,3,4\}$، $D_{ear}=\{1,2,3\}$، $D_{eye}=\{1,2,3\}$.



الشكل8. مثال لمجموعة مختلفة من نماذج الأنف ورأس الذقن والأذن وحجرة العين حيث كل نموذج أسندت إليه قيمة موصف خاصة.

حفظنا في قاعدة المعطيات لكل صورة بنية معطيات للنموذج ثلاثي الأبعاد، والتي تتضمن إحداثيات العقد وزاوية ميلها، و في المناطق حيث لا يوجد تغير كبير في زاوية التوجه، نحفظ في قاعدة المعطيات فقط عقد الحافات وإحداثياتها لتخفيض الحيز اللازم.

ومن أجل الحصول على تمثيل واقعي وصحيح لسطح الوجه، وللاستخدام الفعال لطريقة ع ن م، وجدنا أنه يجب تخزين لكل عقدة فرق زوايا الميل بالنسبة لاتجاهي التقوس الرئيسيين: $\alpha_{max} = \alpha 1 - \alpha 2$, $\alpha_{min} = \alpha 1 - \alpha 3$

تضمنت مجموعة صور الاختبار *15* صورة وجه، عولجت معالجة أولية وخفضت إلى قياس *200x200* بيكسل و *16* مليون تدرج رماديات. الصور في شروط أخذ مختلفة، وزوايا استدارة وجه مختلفة، وشروط إضاءة، و تعابير وجه مختلفة.

في البداية في بعض العينات، لم نستطع استخلاص مميزات الوجه، في عينات أخرى لم يكن استخلاص المميزات صحيحاً تماماً، لعل ذلك بسبب الإضاءة غير المنتظمة، مما قادنا إلى تسوية النصوع في عينات الصور وتدريج الصور إلى القياس نفسه.

بعد عملية المعالجة الأولية لاحظنا تحسيناً ملحوظاً في تحديد المميزات ومن ثَمَّ في الأداء الكلي للنظام. ومن تقييم نتائج التجارب توصلنا إلى ما يأتي:

– نظراً لأنَّنا سنخزن زوايا الميل لعقد معينة للنموذج ثلاثي الأبعاد للوجه في قاعدة المعطيات البيانية، سيُخفض حجم المعطيات على الأقل إلى 33% (مثلاً. من 43 ك بايت إلى 29 ك بايت لكل نموذج)، ويمكن أن يكون ذلك أكثر اعتماداً على التطبيق.

– تسمح صيغة ع ن م لنا بتحقيق التعرف ضمن تركيبة العتاد (مثلاً دارات FPGA) وإرسال المعطيات بسهولة وفعالية عبر وسيط الاتصال بالإضافة إلى مزامنة المعطيات بين المرسل والمستقبل دون الحاجة لإرسال معلومات الساعة.

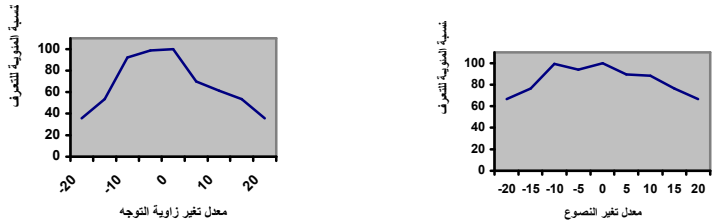
– لأجل التقييم أخذنا نصوعاً معيارياً *Bn* و توجهاً معيارياً (ممثلاً بالزوايا الثلاث *A=B=C=0*) محسوبة لوجه وسطي، وهو وسطي *N* وجهاً.

بينت التجارب التي قمنا بها أن ازدياد النصوع يقلل نسبة التعرف وبأن تغيير اتجاه الوجه إلى الأعلى يقلل نسبة التعرف أيضاً حسب ما هو مبين في الشكل 9.

– تختلف متطلبات تمثيل النماذج ثلاثية الأبعاد عن متطلبات التمثيل من أجل الرؤية الواقعية. بعد بناء النموذج ثلاثي الأبعاد أي عند الانتقال من الفضاء الثنائي إلى الفضاء ثلاثي الأبعاد، من الضروري إيجاد طريقة للرجوع إلى الفضاء ثنائي البعد أو تمثيل المعطيات. عموماً يمكن القول: إنَّ اختيار الطريقة يعود إلى هدف وإستراتيجية حل المهمة.

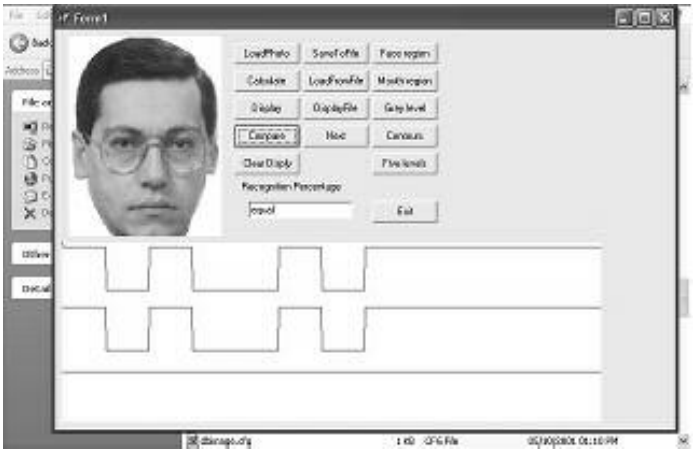
ومن ثَمَّ فان كل ما توصلنا إليه أعلاه يبرر اختيارنا لطريقة التمثيل ع ن م ويجعلهـا طريقة فعالةً. كما نستنتج أنه مع أن الطريقة حساسة لتغيرات الإضاءة و التوجه، لكنها اقتصادية في حفظ المعطيات، والتحقيق العتادي، ومناسبة للتزامن في الاتصالات.

ومن ثَمَّ، للتغلب على مساوئ هذه الطريقة نحتاج لتوحيد شروط الإضاءة و التوجه بين الصور، وأخذ شرائط استخراج الأعمدة على العقد في مواقع محددة ترتبط بنقاط علام الوجه المستخرجة. هذه الإجراءات تزيد التعرف.

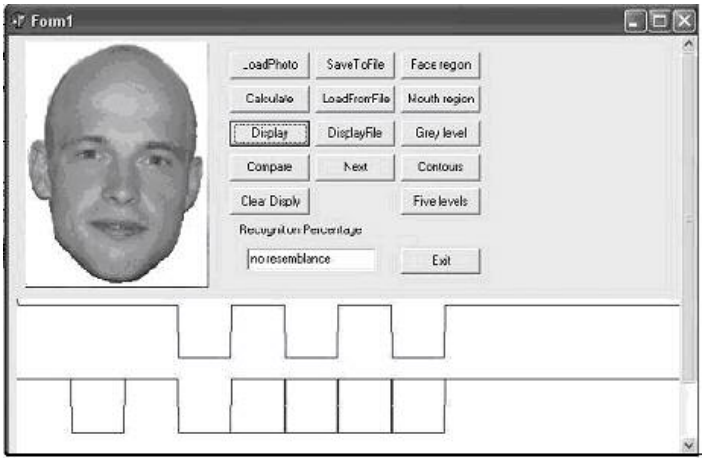


الشكل 9. نسبة التعرف نسبةً إلى معدل تغير النصوع ومعدل زاوية التوجه.

نماذج من شاشات البرنامج مبينة في الشكلين 10 و 11.



الشكل10. عينة لإحدى شاشات البرنامج تبين نتيجة المقارنة لصورة الشخص نفسه باستخدام تمثيل ع ن م، ومبيناً عليها نسبة التعرف equal.



الشكل11. عينة لأحد شاشات البرنامج تبين نتيجة المقارنة لصورة شخصين مختلفين باستخدام تمثيل ع ن م، ومبيناً عليها نسبة التعرف no resemblance.

6. خاتمة:

مع أن كثيراً من طرائق التعرف على الوجوه قد اقترحت منذ عهد بعيد، لكنها لا تزال تعاني من التعقيد والتحقيق المكلف. تعدُّ طرائق بـــروكس [15] و لـــوو [17] غرضـــية التمركز الأكثر دقةً وإيجازاً في وصف الأشكال لكنها تخفق عندما الأغراض بحالـــة تغطية جزئية. يستخدم ديكسون، بنتلاند [16]، و لوو [17] إجراء مقارنة والـــذي يتبنـــى الإسقاط للأجسام ثلاثية الأبعاد على صورة ثنائية الأبعاد، وهي مهمة أصـــعب مـــن الإجراءات الثنائية/ثنائية الأبعاد أو الثلاثية/ثلاثية الأبعاد. طرائق منسكي [18]، برنـــز، رايزمن [19]، بتيجان، بونس، كريغمن [20]، أولمن [23]، برويل [21]، كامبس، شـــابيرو ، هاراليك [22]، ادلمن، و بلتهوف [24] مشهدية التمركز أبسط لكنها تتطلب حيـــز قاعـــدة معطيات أكبر. يتعرف كل من ديكنسن، بنتلاند، وروزنفيلد [25] أولاً على أجزاء عامة بواسطة مشاهدهم المميزة، من ثم تُرَكَّبُ هذه الأجزاء ضمن توصيف جسمي التمركز

للمقارنة مع نماذج أجسام. يستخدم كل من زانغ، سوليفان، و باقر [26] كلا التمركزين المشهدي والجسمي لكل جسم من أجل التعرف. استخدمت التوصيفات مشهدية التمركز لإيجاد مطابقة تقريبية سريعة، من ثم يتم التحقق من هذه المطابقة باستخدام طريقة التمركز الجسمي.

تبين هذه الورقة أن نظامنا من أساس بنيته ثابت ضد كل أنواع التغيرات، ويجمع أفضل المزايا للطرائق الموجودة. حسبنا فروق الزوايا بين نقاط مأخوذة على سطح الوجه وبناءً على هذه الزوايا، تنقل بعدها النقاط إلى فضاء اقليدي منخفض الأبعاد، مما يعطي نموذجَ سطحٍ ثابتٍ متساوي القياس، والذي هو بدوره متماسك ضد التغيرات في الموقع، والإضاءة، وتعابير الوجه. في النهاية يقارن النموذج مع قاعدة معطيات من النماذج.

المراجع

[1] Al-Akkad M. A. and Kochuhanov V. N., A framework for building a general 3D model, Izhevsk, May 2003.
[2] Al-Akkad M. A. and AboulNour H., Design and Implementation of a developed bit-synchronizer for synchronous data communication and its computer simulation for testing and development purposes//Engineering Sciences Journal. – Damascus University, 1996. – In 12 v. – V.2. Thesis Abstract p. 178.
[3] Al-Akkad M. A., AboulNour H., Using curvature and neural networks for recognizing faces//accepted to be published in Damascus University journal.
[4] Mclvor A. M. and Waltenberg P. T., Recognition of simple curved surfaces from 3D surface data. IRL Report 690, Industrial Research Limited, March 1997.
[5] Kumar P., Simplification and Reconstruction ,University of New York at Stony Brook, 1999.
[6] Zhou L. and Kambhamettu C., Extending Superquadrics with Exponent Functions: Modeling and Reconstruction, University of Delaware, 2000.
[7] DeCarlo D. and Metaxas D., Blended deformable models, IEEE Trans. Pattern Anal. Mach. Intell. 18(4), 443–448, 1996.
[8] Fejes S. D. and Rosenfeld A., Discrete Active Models in Vision Geometry, University of Maryland, 1996.
[9] Vetter T. and Poggio T., Image synthesis from a single example image. In Proceedings of the European Conference on Computer Vision, Cambridge UK, 1996.
[10] Kumar P., Simplification and Reconstruction, University of New York at Stony Brook, 1999.
[11] Oosterom. P. J. M. van, Stoter J. E., Zlatanova S., Quak W. C., The balance between Geometry and Topology, Advances in Spatial Data Handling, 10th International Symposium on Spatial Data Handling, Richardson D. and Oosterom P. van (Eds.), Springer-Verlag, Berlin, pp. 209-224, 2002.

[12] Shi W. Z., Yang B. S. and Li Q. Q., An Object-Oriented Data Model for Complex Objects in Three-Dimensional Geographic Information Systems. International Journal of Geographic Information Science, 2002.
[13] Pilouk M., Integrated modelling for 3D GIS, PhD thesis, ITC, The Netherlands, 1996.
[14] Billen R., Zlatanova S., Mathonet P. and Boniver F., The Dimensional Model: a framework to distinguish spatial relationships, in: Advances in Spatial Data handling, Richardson D., Oosterom P. van (Eds.), Springer, pp. 285-298, 2002.
[15] Brooks R. A., Symbolic reasoning among 3-D models and 2-D images. Artificial Intell., 17: 285–348, 1981.
[16] Dickinson S. J. and Pentland, A. P., A unified approach to the recognition of expected and unexpected geon-based objects. In Applications of Artificial Intell. X: Machine Vision and Robotics, vol. 1708, SPIE, pp. 614–627, 1992.
[17] Lowe D. G., Fitting Parameterized 3-D Models to Images. Tech. Rep. 89–26, Dept. of Computer Science, Univ. of British Columbia, 1989.
[18] Minsky M., A framework for representing knowledge. In The psychology of computer vision, ed. by P.H. Winston, pp. 211–277, 1975.
[19] Burns J. B. and Riseman E. M., Matching complex images to multiple 3D objects using view description networks. In Proc. IEEE Conf. Comput. Vis. Patt. Recogn., pp. 328–334, 1992.
[20] Petitjean S., Ponce J. and Kriegman D. J., Computing exact aspect graphs of curved objects: Algebraic surfaces. Int. J. Comput. Vision 9(3): 231–255, 1992.
[21] Breuel, T. M., Geometric Aspects of Visual Object Recognition. Ph.D. dissertation, Dept. of Brain and Cognitive Sciences, Mass. Inst. Technol., 1992.

[22] Camps O. I., Shapiro L. G. and Haralick R. M., PREMIO: An overview. In Proc. Workshop on Directions in Automated CAD-Based Vision, Maui, Hawaii: IEEE Computer Society Press, pp. 11–21, 1991.
[23] Ullman S., Aligning pictorial descriptions: An approach to object recognition. Cognition 32: 193– 254, 1989.
[24] Edelman S. and Bülthoff H. H., Orientation Dependence in the Recognition of Familiar and Novel Views of Three-Dimensional Objects. Vision Res. 32(12): 2285–2400, 1992.
[25] Dickinson S. J., Pentland A. P. and Rosenfeld A., 3-D shape recovery using distributed aspect matching. IEEE Trans. Patt. Anal. Machine Intell. 14(2): 174–198, 1992.
[26] Zhang S., Sullivan G. and Baker K., Using automatically constructed view-independent relational model in 3D object recognition. In Proc. European Conf. Comput. Vis., pp. 778–786, 1992.

تاريخ ورود البحث إلى مجلة جامعة دمشق 2005/6/23.