

خوارزمية كفاءة للكشف عن العين في صور الوجوه

د. لقمان سفر علي
كلية الهندسة/ جامعة الموصل

حسين محمد حسين
كلية الهندسة/ جامعة الموصل

الخلاصة

تعتبر الكشف عن موقع العين من أهم وأصعب المهام في حقل الابصار الآلي. تستغل كشف العين في التطبيقات الطبية وفي التعامل بين الإنسان والآلة وبالأخص الأشخاص ذوي الاحتياجات الخاصة. في هذا البحث قدمت خوارزمية مبسطة وكفاءة في كشف موقع العين في الصور وإطارات الفيديو بالاعتماد على الخواص اللونية وخواص الإضاءة لمنطقة الوجه للكشف عن منطقة الوجه ومن ثم الكشف عن منطقة العين في حدود منطقة الوجه باعتماد الخواص الهندسية مع خواص الإضاءة والخواص اللونية للتحديد الدقيق لمركز بؤبؤ العين، وقد أظهرت النتائج نجاح الخوارزمية في كشف العين في 93% من الصور المستخدمة في الاختبار وبزمن قصير مما يجعلها ملائمة للتطبيقات السريعة.

الكلمات المفتاحية: إبطار آلي، كشف العين.

An Efficient Algorithm for Eye Detection in Faces Images

Hussein Mohammed Hussein
College of Engineering,
University of Mosul

Dr. Luqman Sufer Ali
College of Engineering,
University of Mosul

Abstract

Human Eye's detection is one of the significant difficulties in computer vision. Eye detection utilized in medical applications and computer control system for people with special needs. In this paper, an efficient algorithm for eye detection is proposed. The algorithm utilized both luminesce and chrominance components to detect face region and later with geometrical characteristics to fine detection for pupil center. The experiments results show that the algorithm is works with 93% of images database used for testing with small computation time means the algorithm is usable in real time applications.

Keywords: Computer Vision, Eye Detection.

1- المقدمة

تعتبر عملية الكشف عن العين في الصور وإطارات الفيديو من الأمور المهمة في مجال الإبصار الآلي. حيث يمكن الاستفادة منها في التعرف على هوية الأشخاص وفي التحكم بالحاسوب حيث يعطي تفاعلاً أكثر مع استخدام الماوس أو لوحة المفاتيح وهي ضرورية بالنسبة للأشخاص ذوي الاحتياجات الخاصة. يعطي العين معلومات مهمة يمكن الاعتماد عليها في معرفة التعابير الوجهية للأشخاص عندما يريد الشخص التعبير عن شيء ما، وهذا يستخدم على نطاق واسع في الإعلانات، مع نظرات العين للأشياء الملفتة، يمكن للمعلنين نشر الإعلانات التي تجذب عدد أكثر من الناس ومعرفة أي أنواع الإعلانات تجعل الناس يقررون الشراء من عدمه. كما يستخدم في أنظمة كشف حالة سائقي السيارات لتجنب وقوع الحوادث وتستغل أيضاً في التطبيقات الطبية.

نشرت العديد من الدراسات في مجال كشف العين، وقد أعتمد الباحثون مختلف الطرق للوصول إلى نظام لكشف العين تتميز بالسرعة والمتانة وتعمل بصورة صحيحة في مختلف الظروف ولا تتأثر بوضعية الرأس وخلفية الصورة. ويمكن تقسيم الطرق المتبعة إلى:

- 1- الطرق المعتمدة على المقارنة مع قالب العين.
- 2- الطرق المعتمدة على تمييز المظهر والحواف.
- 3- الطرق المعتمدة على السمات اللونية وخواص الإضاءة والأبعاد الهندسية.

قدم الباحث هونغ من جامعة بكين الصينية في [1] طريقة هجينة تعتمد على كشف الوجه ومن ثم المقارنة مع قالب العين لاستخلاص منطقة العين. وقدم يولي وهانلن من جامعة هارفارد طريقة استخدام القالب المرن في استخلاص السمات في الوجه كالعين والفم، وكانت نتيجتها الدقة في التحديد لكنها تحتاج في البداية إلى تحديد موقع القالب قريباً من المنطقة المراد تحديدها [2].

ومن الطرق المعتمدة على المظهر حيث قام مجموعة من الباحثين بتمثيل حواف العين بـ 16 معلماً، يمكن تمييزها باستخدام مصنف (Classifier) يدرّب على التمييز من خلال المظهر المحيط بالمعالم الستة عشر [3] وهي بذلك قد مكنت من تمييز العين حتى في حالة الإغماض. وفي [4] استخدم الباحث طريقة Harris للكشف عن زوايا العين واعتمدها في استخلاص منطقة العينين، وقد أعطت نتيجة صحيحة في 90% من الصور المستخدمة في الاختبار. ومن الطرق الأخرى والمعتمدة على المظهر وتوجيه أشعة IR للعين، حيث يمكن استقبال الأشعة المنعكسة لتحديد منطقة العين بصورة دقيقة [5] و [6].

وفي عام 2003 قدم مجموعة من الباحثين أسلوب حساب الأبعاد بين المكونات المستخلصة من منطقة الوجه لتكوين مثلث يمثل في رؤوسه العينين ومنطقة الشفاه [7]، ومن الطرق المعتمدة على السمات اللونية والأبعاد الهندسية في [8] و [9] والتي استخدمت في كشف العين ومنطقة الوجه، ومنها أيضاً في [10] وفيها أعتمد الباحث على الخواص اللونية والحواف لاستخراج النقطة المضاءة في العين. وبينك وأخرون قدموا طريقة تجمع ما بين الاعتماد على السمات لتحديد منطقة العين، واستخدام القالب لتحديد مركز الحدقة في [11].

يهدف هذا البحث إلى إيجاد خوارزمية تتسم بالسرعة والكفاءة للكشف عن العين وتحديد مركز سواد العين (البؤبؤ) بالاعتماد على السمات اللونية وخواص الإضاءة للصورة؛ حيث يمكن استخدامها في كشف اتجاه تحديق العين، تستثمر في التحكم بالحاسوب من خلال النظرات، فعند النظر إلى جزء معين من شاشة الحاسوب فإن مؤشر الماوس تنتقل إلى ذلك الجزء.

2- مراجعة لبعض تقنيات معالجة وتحليل الصور 1-2 نماذج الألوان

نماذج الألوان -كغيره من النماذج الرياضية للظواهر الفيزيائية- يمكن تمثيلها بعدة صيغ، لكل منها محاسنها ومساوئها، وكل منها تصلح في تطبيق معين دون غيره. أقترح العديد من نماذج الألوان (Colour Models)، وأغلب النماذج تتكون من ثلاث مركبات للتعبير عن معلومة معينة كأن يكون لون أو إضاءة.

1-1-2 نموذج RGB

في بداية القرن التاسع عشر، قام الفيزيائي توماس يونك واعتماداً على تجاربه باقتراح نموذج ثلاثي الألوان للعين البشري. افترض يونك بوجود ثلاث أنواع فقط من مستقبلات الضوء في العين، واحدة للترددات الواطنة من الطيف المرئي وواحدة للترددات المتوسطة والأخرى للترددات العالية [12]. والجهاز العصبي المركزي للإنسان يتحسس للفرق بين شدة الاضاءة لمختلف أنواع المخاريط [13]. وبعتماد نموذج يونك اقترح نموذج RGB وهو الأكثر شيوعاً حيث أنها تستخدم في كاميرات الفيديو وشاشات الحاسوب. يتمثل النموذج بثلاثة مركبات هي الأحمر والأخضر والأزرق ومنها جاءت التسمية RGB [12]، حيث يتمثل كل لون من ألوان الطبيعة بنسب معينة من الألوان الثلاثة.

2-1-2 نموذج (YCbCr)

وفي أواخر القرن التاسع عشر، قام عالم وظائف الأعضاء هيرنج بعدة تجارب في إدراك العين للألوان واستنتج بأن نظرية يونك لا يفسر جميع النتائج التي توصل إليها بصورة دقيقة، لذا قام بتطوير نظرية جديدة حول إدراك الألوان. اعتقد هيرنج أن أداة الإحساس بالألوان للعين تمتلك ثلاث قنوات مزدوجة، واحدة تحمل المعلومات حول اللون الأسود والأبيض حيث عند التقائهما ينتج الإضاءة أو الظلام. والقناتين الأخرين يحملان المعلومات للون الأحمر والأخضر واللون الأزرق/الأصفر، ولا يوجد ألوان مختلطة [12]، وعلى هذا اعتمد النموذج YCbCr.

حيث أن Y تمثل معلومات عن الإضاءة بالترديج الرمادي، بينما تمثل معلومات الألوان من خلال Cb و Cr. هو الفرق بين اللون الأزرق وقيمة الإضاءة، و Cr هو الفرق بين اللون الأحمر وقيمة الإضاءة. ويمكن الحصول على عناصر نموذج YCbCr من نموذج RGB باعتماد المعادلات الآتية [2]:

$$Y = 0.299R + 0.587G + 0.114B \quad (1)$$

$$Cb = 0.493(B - Y) \quad (2)$$

$$Cr = 0.877(R - Y) \quad (3)$$

2-2 العمليات الشكلية في الصور (Morphological Operations)

العمليات الشكلية هي طرق هندسية للمعالجة غير الخطية للصور وهي أداة مهمة في تحليل الأشكال في الصور الثنائية والصور ذات التدرج الرمادي يعرف العمليات التشكيلية على أنها مزيج من العمليات العددية الأساسية كالجمع والطرح بين الصورة المراد معالجتها A، وصورة صغيرة تمثل شكل صغير B ويسمى بالعنصر الشكلي. إن أهم العمليات الشكلية هما التوسيع (Dilation) والتقليل (Erosion)، إن التوسيع يقوم بإضافة عناصر صورية إلى حواف الأشكال لذا تستخدم في تعظيم الحواف بينما التقليل فإنها تقوم بحذف العناصر الصورية من حواف الأشياء وتستخدم في إزالة الضوضاء والأشياء الصغيرة غير المرغوب فيها.

3-2 عنونة العناصر المترابطة (Connected Component Labeling)

يقصد به تقسيم الصورة إلى أجزاء مكونة من عناصر صورية (pixels) والتي تكون مترابطة فيما بينها، والغرض الأساسي منها هو لاستخلاص المعلومات من الصورة، أو لتهيئتها ليكون كإدخال لنظام ابصار آلي. استخلاص الأشياء من الصورة تعني تعريف العناصر الصورية المكونة لذلك الشيء، لأجل ذلك يعمل مصفوفة بأبعاد مساوية لأبعاد مصفوفة الصورة الأصلية ويعطى عنوان محدد لكل العناصر الصورية العائدة للشيء نفسه [13]، يحدد الأشياء على أساس الترابط بين عناصره الصورية.

يوجد نوعان من الترابط بين العنصر الصوري والعناصر المجاورة لها وهما: (1) الترابط مع أربعة عناصر وفيها يكون العنصران مترابطين إذا كانا متجاورين عمودياً أو أفقياً، (2) الترابط مع ثمانية عناصر وفيها يكون العنصران مترابطين إذا كانا متجاورين عمودياً أو أفقياً أو قطرياً.

يمكن أن يكون العنوان على شكل رقم أو حرف أو لون أو أي رمز آخر، ولا يمكن التعامل مع العنوانين على أساس أنها أرقام ولا يمكن معالجة مصفوفة العنوانين بنفس طرق معالجة الصورة الأصلية. ولأجل فهم كيفية العنونة للعناصر المترابطة نفترض أن لدينا صورة ثنائية، و F هي مجموعة العناصر ذات القيمة واحد في هذه الصورة و F^c هي مجموعة العناصر ذات القيمة صفر في هذه الصورة، والهدف النهائي هي إعطاء عنوان LABEL(p) لكل عنصر صوري في الصورة بحيث أن [14]:

$$1- LABEL(p) = \begin{cases} 0 & \text{if } p \in F^c \\ + & \text{if } p \in F \end{cases} \quad (4)$$

$$2- LABEL(p) = LABEL(q) \Leftrightarrow p \in F \text{ and } q \in F \text{ are in the same connected component.} \quad (5)$$

في البداية تتم تهيئة مصفوفة العناوين كما في المعادلة الآتية:

$$\begin{cases} LABEL^{(0)}(p) = + & p \in F \\ LABEL^{(0)}(p) = 0 & p \in F^c \end{cases} \quad (6)$$

وبعد ذلك وعند المحاولة $t > 0$ ، فإن القيم تتجدد في مصفوفة العناوين باستخدام المعادلات الآتية:

$$LABEL^{(t)}(p) = \min\{LABEL^{(t-1)}(p); \lambda; L_{min}(p)\} \quad (7)$$

حيث أن:

$$L_{min}(p) = \min_{k, l} \{LABEL^{(t-1)}(q) | q = (x_p + k, y_p + l) \in F\} \quad (8)$$

حيث أن λ تمثل عدد الأشياء أو العناصر المترابطة وتزداد قيمته بمقدار 1 كلما يتم اسناد قيمة إلى $LABEL^{(t)}(p)$. المؤشر k والمؤشر l تعرف العناصر المتجاورة للعنصر p ، وبعبارة أخرى فإنه إذا أمكن تغيير العنوان $LABEL(p)$ عند العنصر p إلى القيمة $LABEL(q)$ فإن عنوان العنصر المتجاور يكون $q = (x_p + k, y_p + l)$ ، أما في حالة كون ذلك الاسناد غير ممكناً والقيمة $LABEL(p)$ غير معروفة، يعتبر العنصر الصوري p عائداً إلى شيء آخر ويسند إلى القيمة λ . يعاد العملية عدة مرات للحصول على مصفوفة لا يتغير عناصرها خلال إعادة العملية مرة أخرى وفي النهاية فإن العناصر العائدة إلى الخلفية تحمل العنوان $LABEL(p) = 0$ ، والعناصر العائدة إلى أحد الأشياء تحمل نفس العنوان $LABEL(p) > 0$.

3- خوارزمية كشف العين في الصور وإطارات الفيديو

تفترض الخوارزمية أن تحقق الصورة المدخلة الشرطين الآتيين: (1) أن تكون الصورة محتوية على وجه شخص واحد، (2) وأن لا تحتوي الخلفية على ألوان براقية كالأصفر والأحمر بإضاءة عالية.

يمكن تقسيم هذه الخوارزمية إلى مرحلتين: الأولى مرحلة كشف منطقة الوجه في الصورة، والثانية مرحلة الكشف عن العينين ضمن منطقة الوجه. وقد تم الاستفادة من الخواص اللونية وخواص الإضاءة لأجل ذلك، حيث يعطي نتائج مضمونة وبزمن قصير، لكونها لا تحتاج إلى عمليات رياضية معقدة، الشكل 1 يوضح المخطط الانسيابي للخوارزمية.

1-3 كشف منطقة الوجه

تستخدم الجزء Cr في النموذج اللوني YCbCr حيث أنها تمثل فرق اللون الأحمر عن قيمة الإضاءة، ولأن منطقة الوجه تحتوي على نسبة عالية من اللون الأحمر، فإن استخدام حد عتبة مناسب لتحويل الصورة Cr إلى صورة ثنائية يعطي منطقة الوجه مع بعض المناطق الصغيرة والعائدة إلى خلفية الصورة وما فيها من أشكال، الشكل 2 يوضح عدة نماذج لذلك. تعتمد الخوارزمية بطريقة أوتسو [15] في اختيار حد العتبة حيث أنها تأخذ بنظر الاعتبار عدد العناصر الصورية عند كل مستوى (Histogram). تتم تصفية المناطق الأخرى والمناطق الصغيرة باستخدام عمليات الشكلية وافتراض أن منطقة الوجه هي أكبر منطقة من بين المناطق الموجودة وكون نسبة الارتفاع إلى العرض هي أقل من 1.8 (الشكل 3) تتم استخلاص منطقة الوجه بشكل صحيح (وتم اعتماد هذا لكونها أعطت نتائج دقيقة مقارنة بأعداد ترواح ما بين 1,4 إلى 2,0) وتستخدم عمليات عنونة العناصر المترابطة في سبيل ذلك، وتحدد احداثيات منطقة الوجه بحساب احداثيات أصغر مستطيل يحيط بالمنطقة.

2-3 كشف العينين

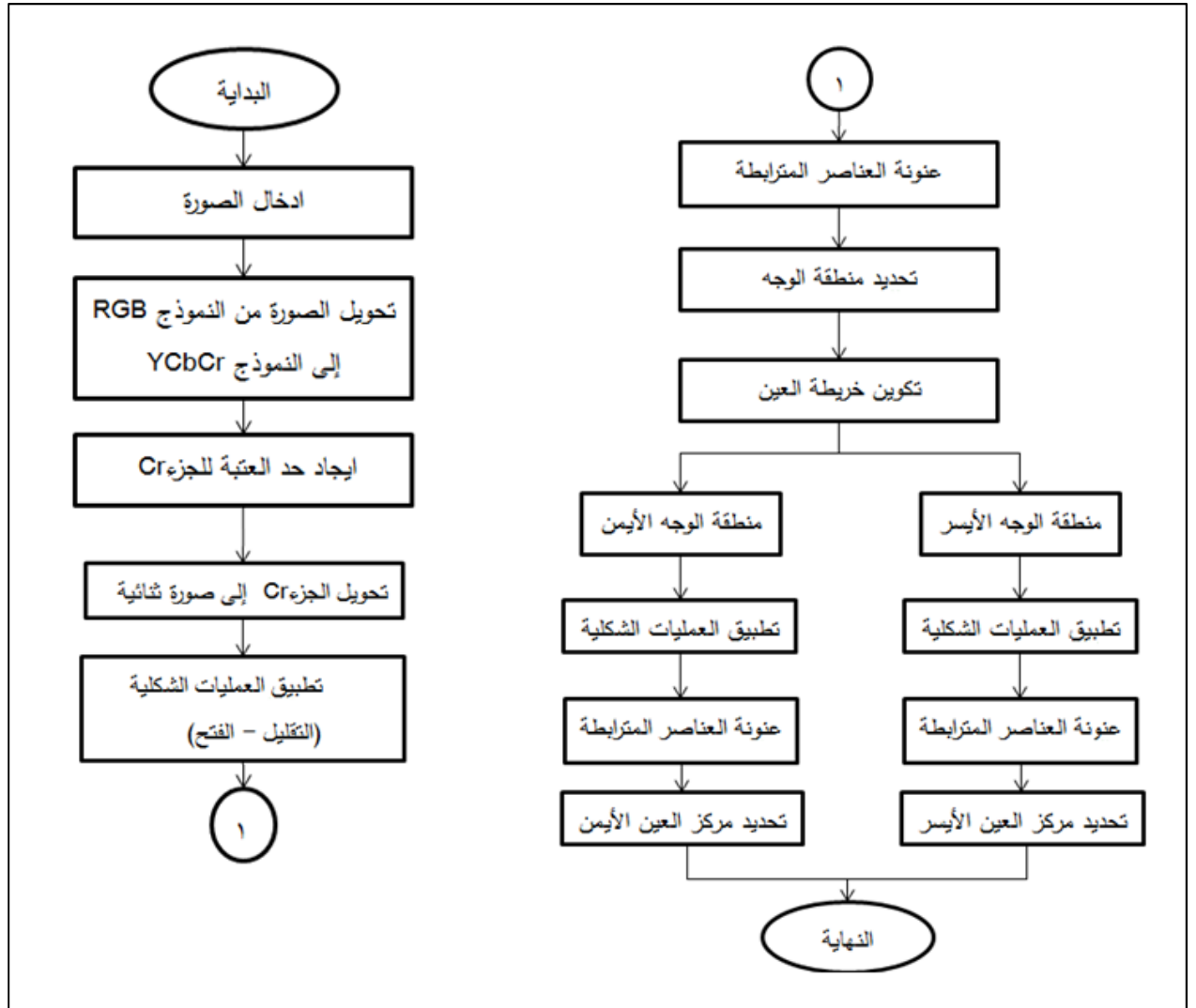
علي: خوارزمية كفاءة للكشف عن العين في صور الوجوه

بعد عملية استخلاص الوجه فإن منطقة العينين أصبحت واضحة بكونها تقع في الجزء العلوي من منطقة الوجه. يمكن تقسيم منطقة الوجه إلى نصفين وبذلك تعرف العين اليمنى واليسرى اعتماداً على الأبعاد الهندسية للوجه. ومن أجل التحديد يستخدم الصورة الثنائية Cr مرة أخرى، ولكن هذه المرة يؤخذ الصورة السالبة لها وضمن منطقة الوجه فقط، ما يعني أن الحسابات تنحصر في جزء معين من الصورة وهي بذلك تساهم في سرعة اتمام المهمة. تبدو منطقة العين واضحة تماماً مع العديد من المناطق الصغيرة المتصلة والتي تمثل عن الحاجبين وأجزاء من الشعر يمكن ترشيح أغلبها باستخدام العمليات الشكلية. ولكن الغاية هي كشف مركز بؤبؤ العين، من أجل ذلك تستخدم الجزء Y من النموذج اللوني YCbCr وبصورة ثنائية وسالبة أيضاً، وقد لوحظ أن استخدام حد العتبة 0.7 لتحويل الجزء Y إلى صورة ثنائية مناسبة تماماً لاستخلاص منطقة البؤبؤ. تجرى عملية ضرب الصورتين الثنائيتين Y و Cr لإيجاد خريطة العين حيث أن:

$$EyeMap(x, y) = \begin{cases} Cr'(x, y) \text{ AND } Y'(x, y) & \text{all } x, y \text{ FaceRegion} \\ 0 & \text{elsewhere.} \end{cases} \quad (9)$$

وتستخدم النسبة 1.2 لاستخلاص منطقة البؤبؤ وتؤشر بأصغر مستطيل يحيط به، وتتم حساب المركز باستخدام المعادلة الآتية:

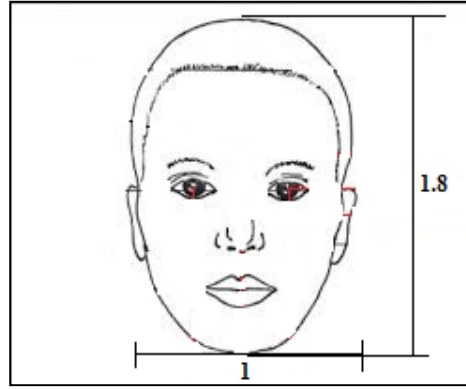
$$Center = \left(\frac{x_{max} + x_{min}}{2}, \frac{y_{max} + y_{min}}{2} \right) \quad (10)$$



شكل 1 المخطط الانسيابي لخوارزمية كشف العين في الصور وإطارات الفيديو.



شكل 2 نماذج لصورة CT بعد التحويل إلى صورة ثنائية.



شكل 3 نسبة الارتفاع إلى العرض لمنطقة الوجه.

4- نتائج تنفيذ الخوارزمية

تم تنفيذ الخوارزمية باستخدام الدوال الموجودة في أداة معالجة الصور في برنامج Matlab، وتم استخدام قاعدة بيانات مكونة من 114 صورة بدقة 640×480 بكسل، وقد استخدمت صورة أمامية وصورة جانبية لاختبار الخوارزمية



شكل 4 نماذج من الصور المستخدمة.

لكل شخص. وقد تبين نجاح الخوارزمية في الكشف الدقيق لمركز بؤبؤ العين في 74% من الصور المستخدمة بالنسبة لكشف كلا العينين، الشكل 4 يبين بعض الأمثلة، بينما نتائج كشف عين واحدة فقد كانت صحيحة في 93% من الصور، في حين أن 7% من الصور المستخدمة لم تفلح فيها الخوارزمية في كشف منطقة العين والسبب في ذلك يعود إلى عدم التحديد الدقيق لمنطقة الوجه.

كما أظهرت النتائج أن وجود النظارات لا تؤثر على أداء الخوارزمية، كما يمكن الكشف عن العينين في حالة كون دوران الوجه بمقدار 20 درجة عن وسط الصورة،

وكما أمكن كشف عين واحدة في حالة كون دوران الوجه بمقدار 45%. إن نسبة نجاح الخوارزمية في كشف العين كافية تماماً بالنسبة إلى الهدف المنشود من البحث، حيث أن وجه الشخص المتعامل مع الحاسوب يكون عادةً بالوضعية الأمامية كما تكون وجه الشخص واضحاً ما يقلل من فرص فشل الخوارزمية في تحديد العين.

5- الاستنتاجات

قدم هذا البحث طريقة جديدة للكشف عن موقع العين في الصور وإطارات الفيديو بالاعتماد على الخواص اللونية وخواص الإضاءة. تتسم هذه الطريقة بالسهولة في التنفيذ وسرعة الأداء. يعطي الخوارزمية مركز منطقة البؤبؤ حيث يمكن استخدامها كإدخال لخوارزمية كشف اتجاه النظر، لذا فإن من الأعمال المستقبلية والتي يمكن الاستفادة من خوارزمية كشف العين فيها هي عمل خوارزمية لكشف اتجاه النظر.

6- المصادر

- 1- Hong Liu, Qing Liu, "Robust Real-Time Eye Detection and Tracking for Rotated Facial Images under Complex Conditions", In: Peking University, Shenzhen Graduate School, Beijing, P.R. China, 2009.
- 2- Alan L. YUILLE, Peter W. HALLINAN, and David S. Cohen, "Feature Extraction from Faces Using Deformable Templates", International Journal of Computer Vision, 8:2, 99-111 (1992).
- 3- Liting Wang, Xiaoqing Ding, Chi Fang, Changsong Liu, Kongqiao Wang, "Eye Blink Detection Based on Eye Contour Extraction", SPIE-IS&T/ Vol. 7245 72450R-1, 2009.
- 4- Dibyendu Ghoshal, Alak Das, "A Study on Human Eye Detection in a Color Image Based on Harris Corner Detector and Sobel Edge Operator", International Journal of Electronics Communication and Computer Technology (IJECCT) Volume 2 Issue 3 (May 2012).
- 5- Antonio Haro, Myron Flickner, Irfan Essa, "Detecting and Tracking Eyes By Using Their Physiological Properties, Dynamics, and Appearance", Appears in Proceedings of CVPR 2000 pp. 163-168.
- 6- Zhiwei Zhu, Qiang Ji, "Robust real-time eye detection and tracking under variable lighting conditions and various face orientations", Computer Vision and Image Understanding 98 (2005), Elsevier Ltd. 124-154.
- 7- Cheng-Chin Chiang, Wen-Kai Tai, Mau-Tsuen Yang, Yi-Ting Huang, Chi-Jaung Huang, "A novel method for detecting lips, eyes and faces in real time", Real-Time Imaging 9 (2003), Elsevier Ltd. 277-287.
- 8- Ali Atharifar, and Sedigheh Ghofrani, "Robust Component-based Face Detection Using Color Feature", Proceedings of the World Congress on Engineering 2011 Vol. II WCE 2011, July 6 - 8, 2011, London, U.K..
- 9- Weimin Huan, and Robert Mariani, "Face Detection and Precise Eyes Location", 2000 IEEE.
- 10- Tanmay Rajpathak, Ratnesh Kumar and Eric Schwartz, "Eye Detection Using Morphological and Color Image Processing", 2009 Florida, Florida Conference on Recent Advances in Robotics, FCRAR 2009.
- 11- Luiz Velho, Alejandro C. Frery, and Jonas Gomes, "Image Processing for Computer Graphics and Vision", Springer-Verlag London Limited 2009, 476.
- 12- Francesco Camastra, PhD, and Alessandro Vinciarelli, PhD, "Machine Learning for Audio, Image and Video Analysis: Theory and Applications", Springer-Verlag London Limited 2008, 484.
- 13- Maria Petrou, and Panagiota Bostdogianni, "Image Processing: The Fundamentals", John Wiley and Sons Ltd, 1999, 355.
- 14- Stphane Marchand-Maillet, and Yazid M. Sharaiha, "BINARY DIGITAL IMAGE PROCESSING: A Discrete Approach", ACADEMIC PRESS, 2000, 279.
- 15- Nobuyuki Otsu, "A Threshold Selection Method from Gray-Level Histograms", IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, VOL. SMC-9, NO.1, JANUARY 1979, 62-66.