

دراسة مقارنة بين طريقتي طرح الخلفية ونموذج مزيج غاوص المستخدمين للتخلص من الخلفية في فيديوهاات ملتقطة في الهواء الطلق

الدكتور إياد حاتم*

هبة حليوه**

(تاريخ الإيداع 14 / 9 / 2014. قُبِلَ للنشر في 14 / 4 / 2015)

□ ملخص □

يعد التخلص من الخلفية الخطوة الأولى في أغلب منظومات رؤية الآلة الرامية إلى عزل كائن معين والتعرف عليه واستخلاص المعلومات منه خصوصاً تلك التي تعتمد على تتبع الحركة (كالتعرف على الإنسان أو تتبع السيارات... الخ). وتشكل الفيديوهاات الملتقطة في الهواء الطلق تحدياً أمام التخلص من الخلفية نظراً لحدوث تغيرات غير مرغوب فيها كتأثير الرياح والإضاءة والظروف المناخية وغيرها.

يقدم هذا البحث دراسة مقارنة بين أبسط طريقة للتخلص من الخلفية (طريقة طرح الخلفية) وطريقة نموذج مزيج غاوص التي تعتبر أكثر تلك الطرق شيوعاً. تتمهذه المقارنة على أساس قدرة كل طريقة على التخلص من الخلفية في فيديوهاات ملتقطة في الهواء الطلق وبشكل خاص في حالة وجود وعدم وجود ظل بالإضافة لتحديات أخرى كحركة الأجسام في الخلفية وتأثير الرياح وعدم استقرار الكاميرا. والانطلاق من نتائج تلك المقارنة لاختيار الطريقة المناسبة لتتبع الحركة في ظروف وشروط مختلفة.

الكلمات المفتاحية: التخلص من الخلفية، طرح الخلفية، نموذج مزيج غاوص، معالجة الفيديو.

* أستاذ مساعد - قسم هندسة الميكاترونك - كلية الهندسة الميكانيكية والكهربائية - جامعة تشرين - اللاذقية - سورية.

** قائم بالأعمال - قسم هندسة الميكاترونك - كلية الهندسة الميكانيكية والكهربائية - جامعة تشرين - اللاذقية - سورية.

Comparative Study between background subtraction and Gaussian Mixture Model algorithms for outdoors videos

Dr. Iyad Hatem*
Hiba Hliwa**

(Received 14 / 9 / 2014. Accepted 14 / 4 / 2015)

□ ABSTRACT □

Background subtraction (BS) is the first step of various computer vision application specially those depending on motion tracking such as (car tacking, human recognition...etc.).Indeed, videos captured outdoors may contain a lot of undesirable changes 'wind impact, illumination changes, weather conditions and others ', generate numerous false positives.

This paper presents comparison between the simplest method for background extraction (background subtraction) and Gaussian Mixture Model which is common method in outdoors videos. These two method are then compared based on the ability of each one to detect moving object in outdoors videos especially with presence and absence of shadow in addition to other challenges like object movement in background, wind effect and camera instability. The results of this comparison is used to determine the suitable method for each state.

Keywords: Background extraction, background subtraction, Gaussian mixture model, video processing

*Associated Teacher, Department of Mechatronics, Faculty of Mechanical and Electrical Engineering, Tishreen University, Lattakia, Syria,

**Academic Assistant, Department of Mechatronics, Faculty of Mechanical and Electrical Engineering, Tishreen University, Lattakia, Syria,

مقدمة:

تحتوي الحركة الحاصلة في مقاطع الفيديو على كمية هائلة من المعلومات عن طبيعة الجسم المرآب. تعتمد كل الطرق المتبعة لرصد الحركة على تتبعها في تسلسل الصور مع الأخذ في الحسبان القيود المفروضة عليها، وتعتمد أول تلك الطرق على أن اختلاف الشدة الضوئية الحاصل في الصور ينتج فقط عن الحركة في الفيديو [1]. ينتج هذا الاختلاف عن حركة البكسلات مع الزمن في الفيديو، أثبتت تلك الطريقة فعاليتها في بيئة متحكم بها من ناحية الإضاءة وثبات الخلفية [2]. لكن كشف الحركة عن طريق طرح الخلفية ليس بالأمر السهل كما يبدو للوهلة الأولى، حيث يسبب التشويش الناتج عن الجودة المنخفضة للكاميرا وعمليات ضغط الصور العديد من الأخطاء في كشف الحركة. ويمكن لهذه الأخطاء أن تنتج أيضاً عن التغيرات غير المرغوبة في الشدة الضوئية كتغيرات الإضاءة الناتجة عن الغيوم والظل والوقت في اليوم أو الظروف المناخية كالمطر والثلج أو حركة النباتات والمياه الناتجة عن الريح أو الأجسام المتحركة في الخلفية [3]. ظهرت طرق أخرى ترتبط بالتغيرات الزمانية والمكانية. حيث تفترض تلك الطرق أن الحركة داخل المشهد لجسم ما تتغير بشكل تدريجي وسلس. أي أن الجسم المتحرك ضمن المشهد لا ينتقل بشكل متقطع ولا يتحرك بشكل مفاجئ بين أطراف الصورة [4]. وقد تم تطوير عدة تقنيات لرصد الحركة الموجودة في المشهد أبرزها طريقتان تستخدم إحداهما تقنية الفروق وتعتمد الأخرى على تقنية مطابقة المنطقة. تستخدم تقنية الفروق التغيرات الزمانية والمكانية للصورة لحساب الحركة الأفقية والشاقولية في سلسلة الصور. أما التقنية المعتمدة على المنطقة فتقوم باقتفاء الحركة الحاصلة في الفيديو عبر مطابقة أجزاء صغيرة في الصورة من إطار إلى آخر [3].

أدى تنوع طرق التخلص من الخلفية إلى ظهور العديد من دراسات المقارنة. احتوت بعض تلك الدراسات على تقييم توصيفي لطرق اكتشاف الحركة [5]. بينما قدمت دراسات أخرى مقارنات كمية تعتمد على تسلسل الحركة في مقاطع الفيديو، كما هو الحال مع توياما [6] وباناها [7] اللذين أجريا دراسات مقارنة لمختلف أنواع خوارزميات التخلص من الخلفية حيث تُستخدم مجموعة وحيدة من البارامترات ثم تُنفذ على مختلف مقاطع الفيديو. تتم بعدها مقارنة العدد الكلي للأخطاء السالبة "الأخطاء الناتجة عن اعتبار أجزاء من الخلفية أنها أجزاء من الكائن" والأخطاء الموجبة "الأخطاء الناتجة عن اعتبار أجزاء من الكائن أنها أجزاء من الخلفية" الناتجة عن عملية التخلص من الخلفية في كل مقطع فيديو معالج. تعد الأخطاء الموجبة والسالبة معيارين نوعيين نموذجيين إلا أنهما غير مستقلين عن بعضهما. فعندما تتناقص الأخطاء السالبة فإن الأخطاء الموجبة تزداد والعكس صحيح. لذلك يعتبر هذين المعيارين غير كافيين معاً لمقارنة طرق طرح الخلفية حيث لا يمكن الجزم بأن انخفاض الأخطاء الموجبة وارتفاع الأخطاء السالبة أفضل أو أسوأ من ارتفاع الأخطاء الموجبة وانخفاض الأخطاء السالبة. علاوة على ذلك فإن الأخطاء الموجبة والسالبة في الدراستين السابقتين نتجت عن تطبيق تعريب محدد مسبقاً لكل طريقة الأمر الذي دعا إلى الاعتقاد أن تبديل العتبات قد يؤدي إلى تحسين الأداء.

قدم هاورد [8] طريقة جديدة في مقارنة طرق طرح الخلفية استناداً إلى ما يسمى بتحليل الاضطرابات. تقوم هذه الطريقة على عدة مراحل، تبدأ المرحلة الأولى بالتخلص من الأخطاء الموجبة في مقطع الفيديو اعتماداً على تعديل قيم التعريب، ثم يتم تعريض الخلفية في سلسلة صور الفيديو إلى شعاع من الاضطرابات في كل اتجاهات الفضاء اللوني RGB، بحيث يحاكي هذا الشعاع حركة الكائن، يُقاس بعدها قدرة خوارزميات طرح الخلفية على كشف الأهداف منخفضة التباين. تعد الميزة الرئيسية لهذه الطريقة هي قدرتها على معالجة كافة أنواع مقاطع الفيديو حتى تلك التي لا تمتلك عزلاً صحيحاً للكائن. هذه الطريقة لا تخلو من العيوب، فعادة ما يكون توزيع البكسلات المنتمية للكائن توزعاً

وحيد النموذج بينما يفترض تحليل الاضطرابات مع خلفيات متعددة النماذج أن الكائن المتحرك هو أيضاً مزيج من التوزيعات متعددة النماذج مما لا يسمح بتقييم الطرق المستندة على المنطقة ولا الاستفادة من أدوات ما بعد المعالجة. تُقيّم بعض الدراسات الاستقصائية الأخرى أساليب طرح الخلفية في سياق الكشف عن الأجسام المتحركة كالسيارات والمشاة وغيرها [9] و [10]. تركز تلك الدراسات على طرق اكتشاف حركة الكائن الذي يتألف من مجموعة من البكسلات المتصلة مع بعضها بعضاً لتشكل كتلة متحركة. يستخدم موقع ومسار تلك الكتلة في عملية التقييم. تكمن المشكلة الأساسية في هذه الطرق من الكشف هو مقدرتها على التعامل مع الكائنات الموصولة جزئياً بعضها مع بعض. حيث تعدها كتلة واحدة كبيرة متحركة. الطريقة الوحيدة لفصل هذه الكائنات هي استخدام مرحلة من عمليات ما بعد المعالجة عالية المستوى. مما يسبب كماً كبيراً من الأخطاء السالبة والموجبة، علاوة على ذلك فقد خصصت بعض الدراسات [10] عمليات ما بعد معالجة خاصة بكل طريقة مما قد يؤدي لزيادة فاعلية إحدى الطرق دون الأخرى وبالتالي الحصول على نتائج المقارنة غير دقيقة.

أهمية البحث وأهدافه:

تكمن أهمية هذا البحث في محاولته الجمع بين بساطة وسهولة طريقة طرح الخلفية وبين استقرار وكفاءة عملية التخلص من الخلفية في الفيديوهات الملتقطة في الهواء الطلق. كما أن عملية المقارنة بين طريقة نموذج مزيج غاوص الذي يعد أكثر الطرق شيوعاً وبين طريقة طرح الخلفية تعطي معلومات هامة عن كفاءة كل منهما والاستخدامات المثلى لهما.

يهدف هذا البحث إلى التخلص من الخلفية باستخدام طريقة نموذج مزيج غاوص بالإضافة لطريقة طرح الخلفية بعد إجراء تعديل بسيط يجعلها أكثر حساسية لتغيرات الإضاءة. تجري بعدها مقارنة الطريقتين السابقتين على فيديوهات ملتقطة في الهواء الطلق في ظروف إضاءة مختلفة، وتتم المقارنة وفق ثلاث مراحل:

1. تقييم الطريقة الأكثر قدرة على التخلص من التشويش
2. تحديد كفاءة كل طريقة في عزل الأجسام المتحركة.
3. تحديد الطريقة المناسبة لكل نوع من الفيديوهات.

طرائق البحث ومواده:

تم التقاط مقاطع الفيديو في الهواء الطلق بواسطة كاميرا بدقة 8 ميغا بكسل بمعدل 18 إطاراً في الثانية حُزنت في ملفات بصيغة mp4 تم تحويلهم بعدها إلى فيديوهات 320X240 بكسل ذات صيغة avi. نُفذت كامل عمليات معالجة الصور واستخلاص المعالم باستخدام برنامج Matlab عبر جهاز حاسوب ذي معالج Intel® Core™ 2Due ضمن نظام تشغيل Windows7 ultimate.

ينقسم العمل في هذا البحث إلى عدة أقسام تبدأ باستعراض خوارزميات طرح الخلفية المستخدمة في هذا البحث وتطبيقه على مجموعة من مقاطع الفيديو الملتقطة لناس تمشي في الشارع بهدف تتبع هذه الأشخاص في الظروف البيئية المختلفة وخصوصاً في حال وجود الظل وعدم وجوده. يتم بعدها اختيار بروتوكول مقارنة للمفاضلة بين الطرق المطبقة وينتهي بعرض للنتائج التي تم الحصول عليها ومناقشتها.

1 التخلص من الخلفية

تعد عملية التخلص من الخلفية في هذا البحث خطوة أولى في تتبع الأجسام المتحركة في مقطع الفيديو. لذلك فإن عملية المقارنة ستتركز على حدود وكتلة الكائن دون غيره وهذا يقتضي القيام بأمرين أساسيين قبل البدء بعملية المقارنة:

1. عزل الجسم المتحرك في مقطع الفيديو باستخدام طرق التخلص من الخلفية.
2. عمليات ما بعد المعالجة والتي تهدف للتخلص من أكبر قدر من الأخطاء الموجبة والسالبة التي لا تؤثر في حدود وحركة الجسم المتحرك في الصورة.

1.1 خوارزميات التخلص من الخلفية:

رغم اختلاف طرق التخلص من الخلفية إلا أن معظمها يشترك في أن تسلسل الصور في مقطع الفيديو يتكون من أجسام تتحرك في خلفية ثابتة علماً أن تلك الأجسام المتحركة مؤلفة من لون واحد أو من (توزع لوني) مختلف عن الخلفية.

1.1.1 طريقة طرح الخلفية (BS) Background Subtraction:

تعد طريقة طرح الخلفية هي أبسط طرق التخلص من الخلفية والتي تعتمد على أن اختلاف الشدة الضوئية الحاصل في الصور ينتج فقط عن الحركة في الفيديو [1] وهذا ما تجسده المعادلة (1).

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + 1) \quad (1)$$

حيث: I الشدة الضوئية في الصورة

δx الإزاحة على المحور x .

δy الإزاحة على المحور y .

تظهر المعادلة السابقة أن التغير في الشدة الضوئية ناتج عن حركة البكسلات مع الزمن في الفيديو. إلا أن نتائج هذه المعادلة تصبح غير صحيحة عند حدوث بعض التغيرات غير المرغوبة في الشدة الضوئية والتي قد تؤثر على الحركة المكتشفة. نذكر من هذه التغيرات:

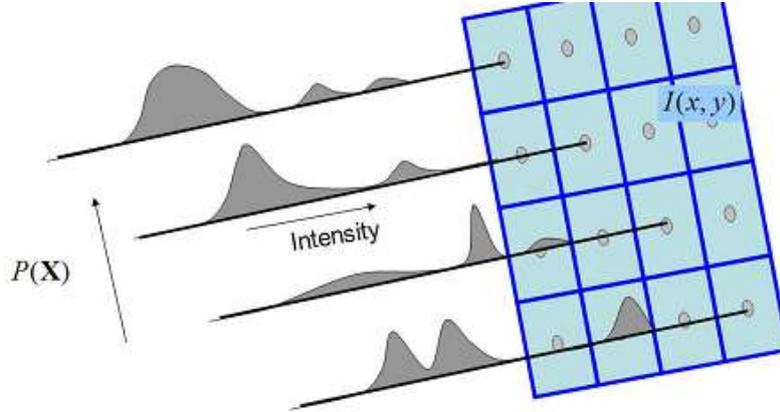
1. التغير في الإضاءة الناتجة عن الغيوم والظل والوقت في اليوم.
2. التغير في الظروف المناخية كالمطر والثلج.
3. حركة النباتات والمياه الناتجة عن الريح.
4. الأجسام المتحركة في الخلفية.

من المعروف أن هذه الطريقة كثيرة الأخطاء ولا تصلح لمعالجة الفيديو الملتقط في الهواء الطلق [3] لذلك تم اقتراح استبدال النظام اللوني RGB بنظام آخر يأخذ التغيرات في الإضاءة في الحسبان وهو نظام HSI اللوني (والذي يتكون من مركبات الصبغة H(Hue) والتشبع S (Saturation) والشدة I (Intensity)).

1.1.2 نموذج مزيج غاوص (GMM) Gaussian Mixture Model:

إن الهدف الأساسي من استخلاص الخلفية هو عزل الأجسام المتحركة عن الخلفية وأي تغيرات قد تحصل فيها. وعندما تُلتقط مقاطع الفيديو في الهواء الطلق تكون الحاجة ملحة لإيجاد نموذج يمثل التباين الطبيعي للشدة الضوئية لكل بكسل في الصورة بما فيها البكسلات الممثلة للأجسام المتحركة كما يظهر في الشكل (1). ويعد نموذج مزيج غاوص من أكثر النماذج المستخدمة في استخلاص الخلفية نظراً لفاعليته واستقراره [11][12].

تُمثل الصورة بمصفوفة من البكسلات. تعبر قيمة كل بكسل عن الشدة الضوئية أو اللون في الصورة. وليكن لدينا المتحول X الذي يعبر عن قيم الشدة الضوئية أو اللون. ولتصنيف بكسلات الخلفية نستخدم نموذج مزيج غاوص



الشكل (1) نموذج يمثل التباين الطبيعي لقيم بعض البكسلات في الصورة [13].

ولإنجاز عملية التخلص من الخلفية يتم إيجاد نموذج للخلفية باستخدام مجموعة من الأطر التي لا تحتوي الأجسام المتحركة. يتم بعدها تحديد تباين قيم البكسلات الموجودة في كل الأطر المختارة لبناء النموذج باستخدام المعادلة الآتية:

$$\sigma_{i,j}^2 = \left(\frac{1}{N-1} \sum_{n=1}^N x_f(i,j) - \bar{x}(i,j) \right)^2 \quad (2)$$

حيث: N هو عدد الأطر المستخدمة لحساب التباين.

n رقم الإطار الحالي.

x_f قيمة البكسل الموجود في النقطة (i,j) في الإطار n .

\bar{x} متوسط قيم البكسلات الموجودة في النقطة (i,j) في كل الأطر المستخدمة في إيجاد التباين.

يستخدم تباين البكسلات لحساب تباين كامل الخلفية وفق المعادلة التالية:

$$\sigma = \sqrt{\frac{1}{o * r} \sum_{i=1}^o \sum_{j=1}^r \sigma_{i,j}^2} \quad (3)$$

حيث o و r هي عدد الأسطر والأعمدة الموجودة في الصورة على التوالي.

ويستخدم الانحراف المعياري لتحديد ما إذا كانت قيمة البكسل في الإطار الجديد المعالج هو جزء من الخلفية.

حيث تُختبر قيم البكسلات المقابلة في الصور الجديدة القادمة وتُحدد الأجزاء التي تختلف قيمها إحصائياً عن الخلفية.

تم تطبيق الاختبار الموجود في [14] لاستخلاص الخلفية وفق ما يلي:

$$P(x_t) = \frac{1}{N} \sum_{i=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_t-x_i)^2}{2\sigma^2}} \quad (4)$$

حيث $P(x_t)$ هو احتمال كون البكسل t جزءاً من الخلفية.

بعد تحديد احتمال كون البكسل في الصورة المعالجة جزءاً من الخلفية يتم تحديد القيمة الاحتمالية لبكسلات

للأجسام المتحركة بحيث تكون هذه القيمة تحت عتبة معينة th . و يمكن تعديل هذه القيمة اعتماداً على معطيات

الفيديو المعالج بحيث تحقق أقل قدر ممكن من الأخطاء، اختُبرت العديد من قيم العتبة (10,30,50,60,70,100) في الفيديوهات المستخدمة في هذا البحث ووجد أن قيمة 60 هي التي حققت أقل قدر ممكن من الأخطاء الموجبة والسالبة ولذلك تم اعتمادها. يظهر الشكل (2) نتيجة تطبيق طريقتي طرح الخلفية ونموذج مزيج غاوص على التوالي على مقطعين مختلفين من الفيديو أحدهما لا يحتوي على ظل وآخر يحتوي على ظل.



إطار لا يحتوي على ظل



طريقة طرح الخلفية



نموذج مزيج غاوص



إطار يحتوي على ظل



طريقة طرح الخلفية



نموذج مزيج غاوص

الشكل (2) طرق التخلص من الخلفية.

1.2 عمليات ما بعد المعالجة:

لا تنتهي عملية التخلص من الخلفية بتطبيق الطرق المذكورة سابقاً. حيث تظهر العديد من الأخطاء السالبة والموجبة في الصورة. تظهر الأخطاء الموجبة على شكل مساحات صغيرة مستقلة تنتمي أصلاً للخلفية وتعتبر خطأً من الجسم المتحرك في حين تظهر الأخطاء السالبة على شكل ثقوب أو مساحات سوداء في الجسم المتحرك والتي تنتج عن احتساب مساحات من الجسم المتحرك أنها من الخلفية. وبما أن الهدف من عملية التخلص من الخلفية هو تتبع الجسم المتحرك فهذا يستدعي إلى القيام ببعض الخطوات للتخلص من هذه الأخطاء دون التأثير في حدود الجسم المتحرك تتلخص فيما يلي:

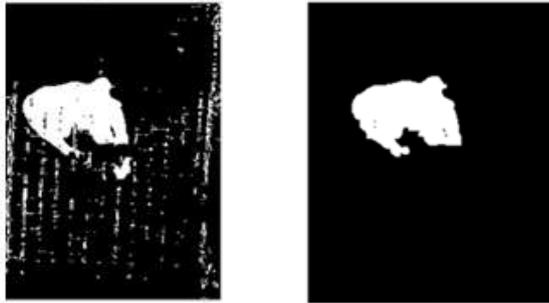
1. التخلص من الأخطاء الموجبة عن طريق الترشيح حسب المساحة والذي يقوم بحساب عدد البكسلات المتصلة مع بعضها بعضاً في أي كائن ثم يحذف كل المساحات التي تقل عن قيمة معينة. اختبرت مجموعة من القيم على عينة من مقاطع الفيديو المعالجة وتم اختيار القيمة التي حققت أقل أخطاء موجبة في كلتا الطريقتين وقدرت هذه القيمة بـ 70.

2. التخلص من الأخطاء السالبة باستخدام عملية التوسيع المورفولوجية والتي تقوم عملية التوسيع بتكبير تدريجي لحواف المناطق في الصورة مما يؤدي إلى ملء الفجوات الصغيرة في الصور الثنائية. ويحدث ذلك عن طريق وضع عنصر البناء "تم تحديدها في هذا البحث بمصفوفة واحدة أبعادها 3×3 " فوق كل بكسل في الصورة، فإذا كان أحد البكسلات الممتلئة للكائن (البكسلات ذات القيمة 1 في هذه الحالة) مقابلة لعنصر البناء هذا عندها يتم تغيير قيم كل البكسلات المقابلة لعنصر البناء لتأخذ قيمة 1.

يبين الشكل(3) نتيجة تطبيق عمليات ما بعد المعالجة المذكورة سابقاً على الصور الناتجة عن التخلص من الخلفية بطريقة طرح الخلفية ونموذج مزيج غاوص على التوالي.



نموذج مزيج غاوص

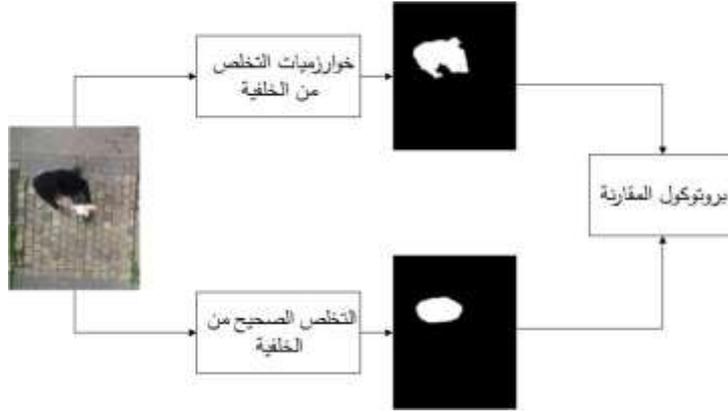


طرح الخلفية

الشكل(3) نتيجة تطبيق عمليات ما بعد المعالجة.

2بروتوكول المقارنة:

لتحديد كفاءة طرق التخلص من الخلفية تتم مقارنة مقاطع الفيديو الناتج عن تطبيق هذه الطرق مع مقاطع فيديو تمثل التخلص الصحيح من الخلفية تم عزل الجسم المتحرك فيها يدوياً. تظهر خوارزمية المقارنة وفق الشكل (4) والتي تم تطبيقها على 25 مقطع فيديو يحتوي 12 مقاطع منها على ظل وبقية المقاطع لا يوجد فيها ظل.



الشكل (4) المخطط الصندوقي لبروتوكول المقارنة

يعتمد هذا البحث في تقييم طرق التخلص من الخلفية على معيارين أساسيين:

2.1 نسبة قمة الإشارة إلى الضجيج:

بما أنه تم اختيار مقاطع الفيديو الملتقطة في الهواء الطلق في تقييم طرق التخلص من الخلفية فمن المهم تحديد مقدرة كل طريقة في التخلص من أكبر قدر من التشويش. لإنجاز هذا التقييم تم اختيار مقياس نسبة قمة الإشارة للضجيج $PSNR[15]$ والذي يتم حسابه وفق المعادلة التالية:

$$PSNR = 10 \log_{10} \left[\frac{(P_N)^2}{MSE} \right] \quad (5)$$

حيث P_N هو أعلى قيمة موجودة في الصورة

MSE هو متوسط مربع الخطأ (Mean Square Error) ويحسب كما يلي:

$$MSE = \frac{\sum_{i=1}^r \sum_{j=1}^o (I(i,j) - M(i,j))^2}{r \times o} \quad (6)$$

حيث I هي الصورة الأصلية و M هي الصورة الممثلة للتخلص الصحيح من الخلفية.

تشير المعادلة (5) إلى أنه كلما زادت قيمة نسبة قمة الإشارة إلى الضجيج كانت النتيجة أفضل.

2.2 مقياس F:

يقوم هذا المقياس [16] بالتعامل مع الصورة على اعتبار أن كل بكسل يتحدد بتصنيف ثنائي بحسب انتمائه للكائن المدروس. ويعبر عن دقة هذا التصنيف بمعاملين هما الاسترداد $recall$ ، والدقة $precision$ ونحصل منهما على مقياس F بإيجاد متوسطهما المتناسق وفق المعادلات الآتية:

$$\frac{\text{عدد البكسلات من الكائن التي تم التعرف عليها بشكل صحيح}}{\text{عدد بكسلات الكائن في النموذج المرجعي}} = \text{الاسترداد}$$

$$\frac{\text{عدد البكسلات من الكائن التي تم التعرف عليها بشكل صحيح}}{\text{عدد البكسلات المصنفة أنها من الكائن المدروس}} = \text{الدقة}$$

عدد البكسلات المصنفة أنها من الكائن المدروس

$$\text{مقياس } F = 2 \cdot \frac{\text{الدقة} \cdot \text{الاسترداد}}{\text{الدقة} + \text{الاسترداد}}$$

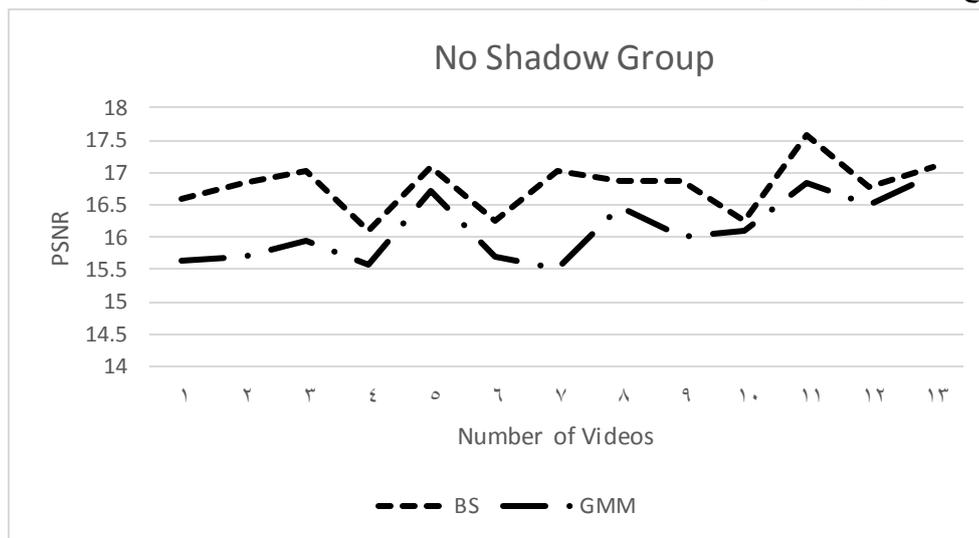
بما أن الاسترداد والدقة يعبران عن صحة التعرف على الكائن نسبة إلى النموذج المرجعي وإلى الكائن المكتشف على التوالي وتتراوح قيمهما بين $[0,1]$ لذلك تتراوح قيمة مقياس F أيضاً بين $[0,1]$ وتكون النتائج أفضل كلما اقتربت من الـ 1

النتائج والمناقشة:

تُسمت مقاطع الفيديو الملتقطة إلى مجموعتين مختلفتين حسب وجود أو عدم وجود ظل وتم اختبار طريقتي التخلص من الخلفية المذكورتين في الفقرة (4) على هاتين المجموعتين ومقارنتهما من حيث كفاءة كل طريقة في التخلص من التشويش وقدرتها على عزل الجسم المتحرك. وبما أن معياري نسبة قمة الإشارة إلى الضجيج ومقياس F تعطيان قيمة لكل إطار معالج فقد تم الحصول على قيمة وحيدة تمثل متوسط القيم التي تم الحصول عليها لكل فيديو معالج.

1-تقييم التخلص من الخلفية في حال عدم وجود ظل:

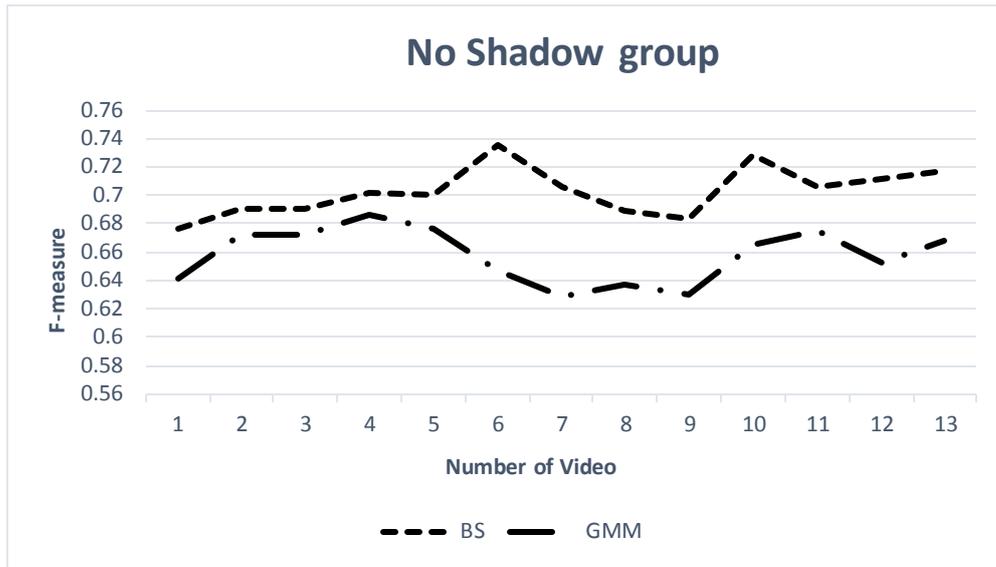
لتقييم كفاءة طريقتي طرح الخلفية ونموذج مزيج غاوص تم اختبارهما على 13 فيديو مختلف لا يحتوي على ظل. يظهر الشكل (5) نتيجة تطبيق معيار نسبة قمة الإشارة إلى الضجيج على مقاطع الفيديو الناتجة عن طرح الخلفية ونموذج مزيج غاوص وبعد تطبيق عمليات ما بعد المعالجة.



الشكل (5) نتائج اختبار نسبة قمة الإشارة إلى الضجيج في حال عدم وجود ظل.

كانت نتائج تطبيق معيار نسبة قمة الإشارة إلى الضجيج مفاجئة فعلى الرغم من القول الشائع [3] بأن طريقة طرح الخلفية تعطي الكثير من الأخطاء وغير صالحة في البيئات غير المتحكم بها إلا أن استخدام هذه الطريقة مع عمليات ما بعد المعالجة التي تم التطرق إليها في الفقرة (4) رفعت من كفاءة هذه الطريقة وجعلتها تتفوق على طريقة نموذج مزيج غاوص كما يظهر في الشكل (5).

أما بالنسبة لمقدرة كل طريقة على عزل الجسم المتحرك في الصورة فيوضح الشكل (6) نتائج تطبيق مقياس F

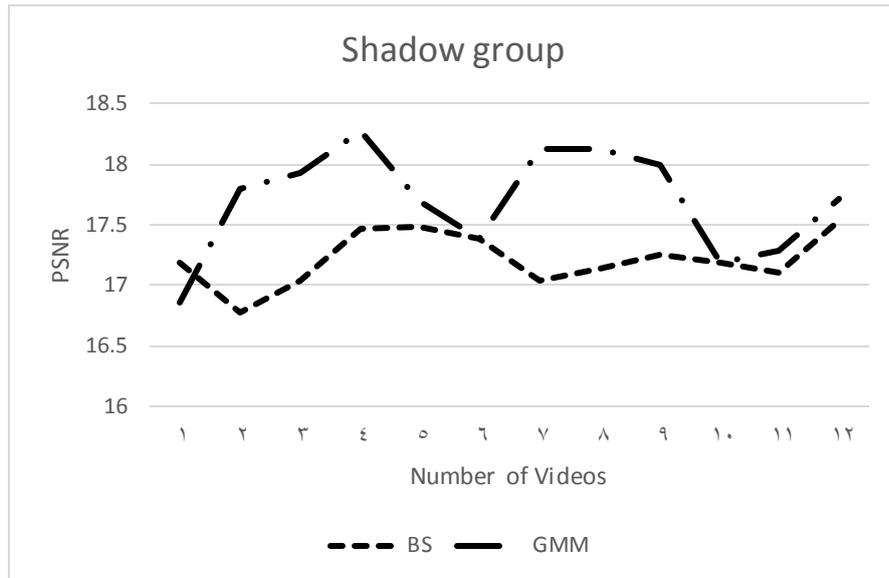


الشكل (6) تطبيق مقياس F في حال عدم وجود ظل

يظهر مقياس F أيضاً كفاءة طريقة طرح الخلفية مع عمليات ما بعد المعالجة في عزل الجسم المتحرك وتفوقها على طريقة نموذج مزيج غاوص في حال عدم وجود ظل. مما يدل على أن الاختيار الصحيح للصورة المرجعية التي تمثل الخلفية بالإضافة لعمليات ما بعد المعالجة التي تستهدف التشويش دون المس بحدود الجسم المتحرك المعزول استطاعت أن تتفوق على طريقة نموذج مزيج غاوص في حال عدم وجود ظل.

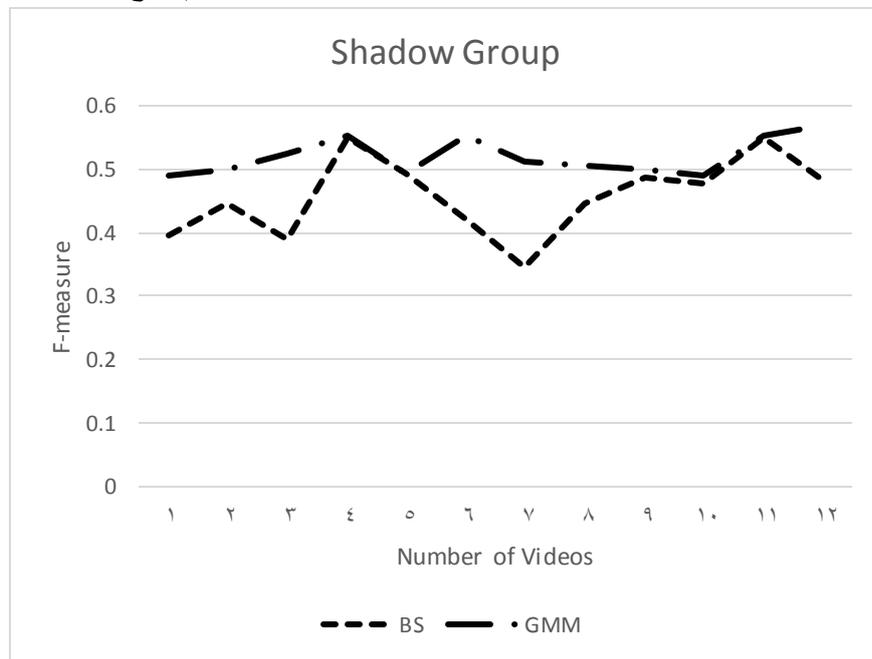
2-تقييم التخلص من الخلفية في حال وجود ظل:

يشكل وجود الظل في مقاطع الفيديو الملتقطة تحدياً كبيراً أمام طرق التخلص من الخلفية. تم اختبار طريقتي طرح الخلفية ونموذج مزيج غاوص على 12 فيديو تحتوي على ظل تتراوح قيم الأطر فيها بين الـ 50 والـ 70 إطاراً يتواجد الكائن المراد تتبعه في 30 إلى 50 منهم. يوضح الشكل (7) نتيجة تطبيق معيار نسبة قمة الإشارة للضجيج على طريقتي طرح الخلفية.



الشكل (7) تطبيق مقياس نسبة قيمة الإشارة إلى الضجيج في حال وجود ظل.

على عكس مقاطع الفيديو التي تحتوي على ظل فإن الشكل (7) يبين أن نموذج مزيج غاوص كان بشكل عام أفضل في التخلص من التشويش في حالة وجود ظل من طريقة طرح الخلفية. أما بالنسبة لمقياس F والذي تظهر نتيجة تطبيقه على مقاطع الفيديو التي تحتوي على ظل في الشكل (8) يثبت نموذج مزيج غاوص أيضاً عبر مقياس F أنه أقدر على التخلص من الخلفية باستخدام طرح الخلفية



الشكل (8) تطبيق مقياس F في حال وجود ظل.

تشير مناطق التقارب في مقياسي PSNR ومقياس F إلى قدرة عمليات ما بعد المعالجة في بعض الأحيان على التخلص من الأخطاء الموجبة والسالبة الموجودة في الفيديو.

الاستنتاجات والتوصيات:

استعرض هذا البحث مقارنة بين طريقة طرح الخلفية ونموذج مزيج غاوص وذلك لأن طريقة طرح الخلفية تعد أبسط طرق التخلص في الخلفية وطريقة نموذج مزيج غاوص والتي تعد من أكثر الطرق الإحصائية شيوعاً. تم اختبار هاتين الطريقتين من حيث قدرتهما على التخلص من التشويش وعزل الجسم المتحرك على مقاطع فيديو ملتقطة في الهواء الطلق وتم تقسيم هذه المقاطع إلى قسمين أحدهما يحتوي على ظل والآخر لا يحتوي على ظل. تم العزل اليدوي لكل فيديو على حده للحصول على فيديو يمثل العزل الصحيح للجسم المتحرك.

تم تقييم طريقتي التخلص من الخلفية باستخدام معيارين أولهما هو معيار نسبة قمة الإشارة إلى الضجيج والتي تختبر مقدرة كل من الطريقتين على التخلص من التشويش في الصورة. أما المعيار الثاني فهو مقياس F الذي يقوم بالتحقق من كفاءة الطريقتين في عزل الجسم المتحرك. ومن النتائج التي تم استعراضها في الفقرة (6) يمكن تلخيص النتائج وفق ما يلي:

1. على عكس ما ذكرته بعض المراجع [17][10] أن طريقة طرح الخلفية أقل كفاءة من طرق التخلص من الخلفية الأخرى في مقاطع الفيديو الملتقطة في الهواء الطلق فإن طريقة طرح الخلفية كما ورد في النتائج المذكورة سابقاً استطاعت التغلب على طريقة نموذج مزيج غاوص في حال عدم وجود ظل. وذلك نظراً إلى أن الدراستين السابقتين اعتمدتا على فيديوهات ملتقطة في الهواء الطلق ذات حركة نسبية كبيرة للخلفية مقارنة بحركة الكائن المراد تتبعه. أما الفيديوهات الموجودة في هذا البحث والتي لا تحتوي على ظل فكانت الحركة النسبية للخلفية مقبولة مقارنة بحركة الكائن واستطاعت عمليات ما بعد المعالجة التخلص منها.
 2. استطاعت عمليات ما بعد المعالجة التقليل من الأخطاء الموجبة والسالبة على حد سواء وبالتالي سببت استقراراً أكبر في مقدرة طريقة طرح الخلفية على التخلص من التشويش في كل مقاطع الفيديو التي تم اختبارها. كما استطاعت هذه العمليات التقريب بين طريقة طرح الخلفية وطريقة نموذج مزيج غاوص في النتائج.
 3. كما ورد في المرجع [17] فإن النتائج المعروضة تثبت أن طريقة نموذج مزيج غاوص لا تعطي دائماً نتائج دقيقة، وخصوصاً في مقاطع الفيديو التي لا تحتوي على مقدار تشويش كبيرة. بينما تتصرف بشكل أفضل عندما تكون الخلفية غير مستقرة أو مع وجود مقدار كبير من التشويش كما في حال وجود ظل.
- استناداً للنتائج المعروضة فلا توجد طريقة رابحة بشكل مطلق في هذه المقارنة. حيث أثبتت طريقة طرح الخلفية كفاءتها في التخلص من الخلفية في حال عدم وجود ظل أو بالأحرى في البيئات المتحكم بها أو تلك التي تكون فيها الحركة النسبية للخلفية صغيرة ويمكن التخلص منها بعمليات ما بعد المعالجة في حين كانت طريقة نموذج مزيج غاوص أكثر كفاءة ودقة في التخلص من الخلفية في حال وجود ظل أو في حال وجود حركة نسبية كبيرة للخلفية مقارنة بالكائن.

المراجع:

- [1] B. a. S. B. Horn, "Determining optical flow.," *Artificial Intelligence*, vol. 17, no. 1, pp. 185-203, 1981.
- [2] Q. Z. a. J. Aggarwal, "Tracking and classifying moving objects from video," in Performance Evaluation of Tracking Systems Workshop, 2001.
- [3] B. A. Smith, Determination of Normal or Abnormal Gait Using a Two Dimensional Video Camera, Blacksburg, Virginia: Polytechnic Institute and State University, 2007.
- [4] E. R. Davies, Machine Vision, 3rd edition, New York: Morgan Kaufmann, 2005.
- [5] T. T. L. W. a. S. M. W. Hu, "A survey on visual surveillance of object motion and behaviors," in Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2004, p. 334:352.
- [6] J. K. B. B. a. B. M. K. Toyama, "Wallflower: principles and practice of background maintenance," in International Conference on Computer Vision, 1999.
- [7] S. S. S. H. a. N. G. S. Panahi, "Evaluation of background subtraction methods," in International conference on Digital Image Computing: Techniques and Applications, 2008.
- [8] K. K. D. H. a. L. D. T.H. Chalidabhongse, "A perturbation method for evaluating background subtraction algorithms," in International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2003.
- [9] B. E. a. C. R. Y. Benezeth, "Comparative study on foreground detection algorithms for human detection," in International Conference on Image and Graphics, 2007.
- [10] J. N. P. R. E. A. P. M. S. P. T. L. R. E. R. F. J. V. a. J. C. D. Hall, "Comparison of target detection algorithms using adaptive background models," in International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2005.
- [11] S. S. C. a. C. Kamath, "Robust techniques for background subtraction in urban traffic video," In Proc. Electronic Imaging: Visual Comm. Image Proc., p. 881–892, 2004.
- [12] D. S. Lee, "Effective Gaussian mixture learning for video background subtraction," *IEEE Trans Pattern Analysis Machine Intelligence*, vol. 27, no. 5, p. 827–832, 2005.
- [13] R. Farnoosh, G. Yari and B. Zarpak, Image Segmentation using Gaussian Mixture Models, Tehran, Iran: Department of Applied Mathematics, University of Science and Technology, 2008.
- [14] D. Gavrilu, "The Visual Analysis of Human Movement: A Survey," *Computer Vision and Understanding*, vol. 73, no. 1, pp. 82-98, 1999.
- [15] J. Boyd, "Synchronization of Oscillations for Machine Perception of Gaits," *Computer Vision and Image Understanding*, vol. 96, no. 1, pp. 35-59, 2004.
- [16] S. Brutzer, B. Hoferlin and G. Heidemann, Evaluation of Background Subtraction Techniques for Video Surveillance, Stuttgart, Germany: Intelligent Systems Group, University of Stuttgart, 2011.
- [17] Y. Benezeth, P. M. Jodoin, B. Emile, H. Laurent and C. Rosenberger, "Comparative Study of Background Subtraction Algorithms," *Journal of Electronic Imaging*, vol. 1, no. 19, 2010.