Using Kernel Estimator and K-Mean Clustering Approach to Hand Gesture Recognition

Ban A. Mitras **AAlA Mahmood Mohammed**

dr.banah.mitras@gmail.com

College of Computer Sciences and Mathematics
University of Mosul, Mosul, Iraq

Received on: 26/09/2012 **Accepted on: 30/01/2013**

ABSTRACT

In this paper we used non parametric density estimator (Kernel Estimator) to estimate probability density function for Image data of Hand Gesture and warping Hand Gesture and we see the curve for Kernel Estimator and combine the curve between Kernel Estimator and normal Distribution. Programs written using the language Matlab (R2009a) .

Keyword: Hand Gesture, Kernel estimator, Pattern Recognition.

استخدام المقدّر أللبي وأسلوب العنقدة بمعدل- k للتعرف على إيماءة اليد

بان احمد متراس **آلاء محمود محمد**

كلية علوم الحاسوب والرياضيات
جامعة الموصل، الموصل، العراق

تاريخ استلام البحث: 2012/09/26 **تاريخ قبول البحث: 2013/01/30**

الملخص

تم في هذا البحث استخدام مقدر الكثافة اللامعلمية (المقدر أللبي Kernel Estimator) لتقدير دالة كثافة الاحتمال لبيانات صورة إيماءة اليد الاعتيادية والمشوهة، وقد تبين أن مقدر الكثافة أللبي قادر على تقدير دالة الكثافة الاحتمالية لصورة إيماءة اليد وتمييزها بواسطة تنعيم وتمهيد بيانات الصورة وجعلها مشابهاً أو مطابقاً لمنحني التوزيع الطبيعي فهذا يعني أنه قد تم تمييز إيماءة اليد، كما تم استخدام تقنية تجزئة الصور (K-Means) للتعرف على إيماءة اليد الاعتيادية والمشوهة. وقد تمت كتابة البرنامج باستخدام لغة(R2009a) Matlab.

الكلمات المفتاحية: ايماءة اليد، المقدر اللبي، تمييز الأنماط.

1. مقدمة Introduction:

أن تعقد الظواهر وتشعبها وتطور العلاقات وعدم توافر معلومات كاملة عنها كأن تكون الظاهرة تحدث لأول مرة أو كون النماذج المفترضة لا تحتوي على خصائص يمكن التعويل عليها، كل هذه الأسباب جعلت الباحث يلجأ إلى استخدام أسلوب يكون أكثر مرونة في التعامل مع البيانات ويتعامل مع طرائق لا تحتاج إلى فرضيات أو في أقل تقدير تحتاج إلى فرضيات أقل لاسيما فيما يتعلق بعدم الحصول على معلمات أو أن بيانات

العينة تكون رتبوية أو نسبوية أو وصفية، وهي موجودة في الكثير من حالات التطبيق ويطلق على هذا الأسلوب بالنمذجة اللامعلمية (Nonparametric Modeling) [8].

من الطرائق المهمة في التمييز هي تمييز الحركات البشرية بوساطة الإيماءة لما لها من أهمية خاصة في لغة الصم والبكم، ويتم أخذها بوصفها إشارة مدخلة. على سبيل المثال نظام تمييز الإيماءة يتم بوساطة متسلسلة لعدد من الأوقات لحركة اليد [10].

إن الافتراض الأساس للأنموذج الإحصائي يتمثل في إمكانية تمييز الإشارة وبصورة جيدة بوصفها عملية عشوائية معلمية، وأن معلمات العملية التصادفية يمكن أن تحدد أو تُقَدّر وبدقة بالنسبة للتطبيقات، وعلى نحوٍ خاص في عملية تمييز الإيماءة، فإنه يمكن تطبيق نماذج الإشارة التصادفية ونماذج الإشارة المحددة وبنجاح.

إن استخدام الطرائق اللامعلمية ظهرت باكرا في بداية القرن العشرين، كما ظهر أن هناك عدداً لا يحصى من الباحثين قاموا بدراسة الطرائق اللامعلمية وطرائق تقديرها، فقد قدم الباحث ذنون في عام 1998 بحثا يستعرض فيه التقدير اللبي بوصفه أسلوبا بيانيا حديثا [2].

وقد قام الباحث حمو في 2000 بمقارنة ثلاثة مقدرات لدالة الانحدار اللامعلمية مقترحا دالة لبية kernel جديدة [1] .

2. المقدر أللبي Kernel Estimator:

هو أسلوب بياني لامعلمي، يمكن استخدامه لتقدير أي دالة إحصائية، أما خصائصه فهي البساطة وسهولة برمجته على الحاسوب ورسم مقدراته بيانياً، كذلك مقدراته تكون مستوية Smooth وغير متحيزة بالمحاذي Unbiased Asymptotically ومتسقة بالمحاذي Consistent Asymptotically ويعمل على تعديل المشاهدات وعلى تقدير دالة الانحدار اللامعلمي التقديرية إلى دالة الانحدار اللامعلمي الحقيقية.[1]

إن مقدر الكثافة أللبي (Kernel) والذي تم اقتراحه من قبل الباحثين Rosenblatt (1956) و(1962) Parzen، يشير إلى نوع عام من أساليب التقدير اللامعلمي للدوال، فإذا فرضنا وجود مجموعة بيانات أحادية المتغير ويراد عرضها بيانيا، فإن المقدر أللبي Kernel سوف يزودنا بوسيلة فعالة لتحقيق هذا الهدف وإيجاد تركيب لمجموعة البيانات دون افتراض الأأنموذج ألمعلمي. [7][1]

لنفرض أن $x_1, x_2, ..., x_n$ هي عينة عشوائية من المتغير العشوائي المستمر X والذي له دالة كثافة احتمال مجهولة $f_x(x)$. إن المقدر أللبي لهذه الدالة هو: [2]

$$\hat{f}_x(x) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x - X_i}{h}) \qquad \ldots(1)$$

إذ إن $h>0$ تمثل معلمة التمهيد Smoothing Parameter أو تسمى سعة القيد Bandwidth ، أما K(.) فهي دالة لُبية (أو دالة النافذة Window Function) وهي عبارة عن دالة رياضية حقيقية محددة تقوم بتسوية المقدر وإزالة التعرجات منه (أي تنقيته، ولهذا سميت نافذة) وتحقق الشرطين الآتيين:

I $K(x) \geq 0 \qquad for\ all \quad x \geq 0$.

II $\int_{-\infty}^{\infty} K(x)\,dx = 1$.

أي إن الدالة اللبية يجب أن تحقق شرطي دالة كثافة الاحتمال كليهما [1].

2.1 اختيار الدوال اللبية Function Kernel Selection:

تستعمل الدوال اللبية Kernel في تقدير كل من دوال الكثافة الاحتمالية، دوال الانحدار ودوال الطيف (Spectral). فهناك سلسلتان من الدوال اللبية Kernel يمكن تمييزهما وهما الدوال اللبية ذات اقل تباين (Minimum Variance Kernel) والتي تعمل على تقليل التباين المحاذي والدوال اللبية المثلى (Optimal Kernel) والتي تعمل على تقليل متوسط مربع الخطأ التكاملي Mean Integrated Square Error (MISE) المحاذي، أي اشتقاق MISE بالنسبة لدالة Kernel .[4] [1]

وإن الدالة اللبية Kernel تكون حقيقية ومتماثلة ومحددة ومستمرة، وتكاملها مساو إلى الواحد [9][7].

2.2 سعة القيد Bandwidth :

المعلمة التمهيدية هي h والتي تدعى bandwidth أي عرض الحزمة، معلمة حجم نافذة الانتشار أو التركيز أو معلمة التباين.

تؤثر المعلمة التمهيدية تأثيرا كبيرا في تمهيد المنحني المقدر واقترابه من المنحنى الحقيقي، إذ بزيادة هذه المعلمة نحصل على تعظيم التحيز وتصغير التباين، وإذا قمنا بتصغير المعلمة التمهيدية يحدث العكس. [5][1]

2.3 التمييز باستخدام مقدر الكثافة اللامعلمية المقدر أللبي (Kernel Estimator):

يتم التمييز باستخدام مقدر الكثافة اللامعلمية (المقدر أللبي Kernel Estimator) وذلك بملاحظة المنحني الناتج من دالة المقدر أللبي، فإذا كان قريباً أو مشابهاً إلى منحني التوزيع الطبيعي فهذا يعني أن المقدر قام بتمييز الصورة الحاوية على إيماءة اليد، أما إذا كان بعيداً عن منحني التوزيع الطبيعي فهذا يعني أنه لم يتم تمييز الصورة الحاوية على إيماءة اليد.[6]

4. تقنية التجزئة بمعدل – k k-Means Technique :

هي إحدى تقنيات العنقدة المكتشفة في سنة 1967 من قبل العالم MacQueen وعدّها تقنية تصنيفات غير مرشدة وتعتمد على معيار اقل مسافة، إذ تعطى مجموعة من القيم ومحاولة تجزأتها إلى K من العناقيد، وتستعمل خوارزمية تكرارية تقلل مجموع المسافات المربعة من الكيان إلى مركز العنقود ولكل العناقيد [6] ولتنفيذ هذه التقنية نطبق الخوارزمية الآتية:

الخطوة الأولى: اختيار مراكز العناقيد عشوائيا $M_1, M_2, \dots, M_k$.

الخطوة الثانية: تنسيب كل نقطة X إلى واحد من العناقيد القريبة لها حسب المسافة بين النقطة ومركز العنقود وعلى أساس المسافة الإقليدية وباستعمال العلاقة الآتية:

$$X \in C_j(t) \quad \text{if} \quad d((X, M_j(t)) < d(X, M_i(t)) \qquad \dots(2)$$

إذ أن $C_j(t)$ يمثل عنقود مركزه $Z_j(t)$ و $i=1,2,\dots$ $j \neq i, j = 1,2,\dots,m$,

d يمثل المسافة الإقليدية بين النقطة X ومركز العنقود M.

الخطوة الثالثة: حساب قيم مراكز عناقيد جديدة $M_j(t+1)$ لكل $j = 1,2,\dots, M$ إذ تصغر المسافات المربعة لكل المتجهات في العنقود $C_j(t)$ بالنسبة لمركز العنقود الجديد وعليه يمكن حساب مركز العنقود الجديد حسب المعادلة الآتية:

$$M_j(t+1) = \frac{1}{S_j} \sum_{X \in C_j(t)} X \quad \ldots(3)$$

إذ أن:

S_j تمثل عدد النقاط التابعة للعنقود $C_j(t)$.

X تمثل نقطة تابعة للعنقود $C_j(t)$.

الخطوة الرابعة: إذا كان $M_j(t+1) = M_j(t)$ $j= 1,2,...,M$ فهذا يعني تقارب الخوارزمية، ثم تتوقف وبخلافه يتم الذهاب إلى الخطوة الثانية.

وهذه الخوارزمية تتغير تبعا للاتي: 1. عدد العناقيد. 2. اختيار القيم البدائية. 3. خصائص البيانات.

5. خوارزمية تجزئة وتقدير دالة كثافة الاحتمال لصورة إيماءة اليد الاعتيادية

Segmentation and Estimator Function Algorithm for Hand Gesture Image

فيما يأتي خطوات الخوارزمية المقترحة:

الخطوة الأولى: قراءة المقطع الفيديو الذي يحتوي على الإيماءة، ثم أخذ إطار منه على شكل صورة وقراءتها.

الخطوة الثانية: تقدير دالة كثافة الاحتمال لبيانات صورة الإيماءة باستخدام المقدر اللامعلمي (المقدر أللبي) .

الخطوة الثالثة: إدخال ضوضاء على الصورة المأخوذة على شكل إطار.

الخطوة الرابعة: إجراء عملية انحلال خطي للصورة.

الخطوة الخامسة: حساب قانون كاوسيان الخطي (قانون الانحلال) . $Y = AX + n$ إذ أن A تمثل مصفوفة القيم المميزة و n نسبة الضوضاء المضافة للصورة.

الخطوة السادسة: إزالة الضوضاء من الصورة باستخدام المرشح الأوسطي.

الخطوة السابعة: أخذ تحدد للحافات باستخدام محدد كاني لإظهار حدود الصورة وحافاتها.

الخطوة الثامنة: إجراء ترشيح للصورة مع المرشح المكون وخزنه في مصفوفة مثل J باستخدام أحد المرشحات canny; Sobel ; log; Laplace,…

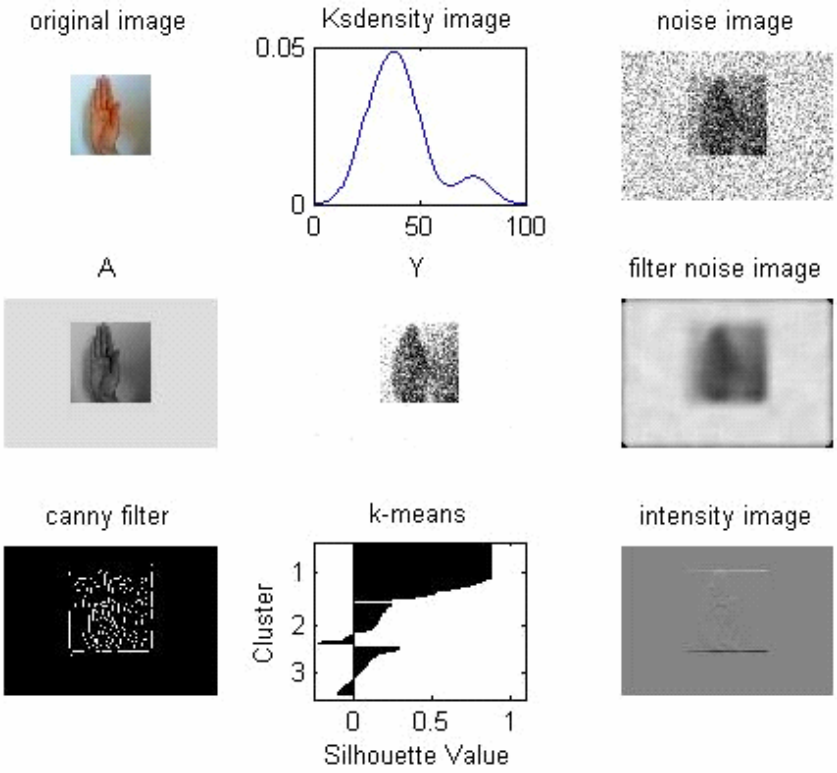
الخطوة التاسعة: إضافة كثافة للصورة المرشحة وخزنها في مصفوفة مثل K

الخطوة العاشرة: إيجاد الوسط الحسابي والتباين للصورة المرشحة K

الخطوة الحادية عشر: إيجاد دالة الكثافة الاحتمالية للتوزيع الطبيعي وخزنه في مصفوفة مثلY

الخطوة الثانية عشر: تجزئة صورة إيماءة اليد باستخدام تقنية K-Means والتعرف عليها.

الخطوة الثالثة عشر: عرض الصور كما في الشكل المرقم (1).



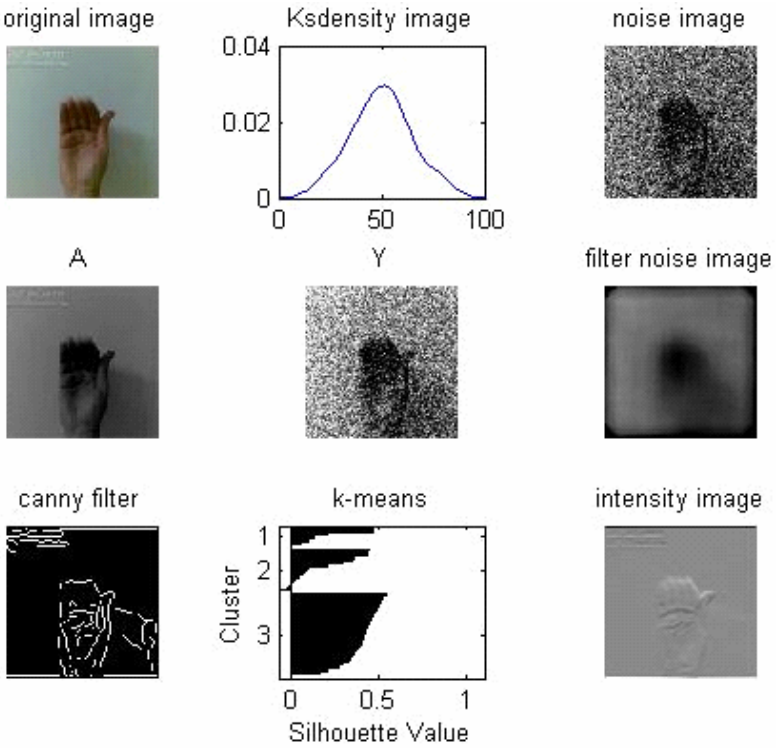
الشكل (1). يمثل الصورة الاعتيادية الأولى التي طبقت عليها الخوارزمية اللامعلمية لتقدير دالة كثافة الاحتمالية لصورة إيماءة اليد الاعتيادية

جدول (1). يمثل مراحل أسلوب العنقدة بمعدل k (k-mean) لصورة إيماءة اليد الاعتيادية في الشكل (1)

iter	phase	num	sum
1	**1**	**210**	**2.13983e+008**
2	**1**	**36**	**2.04656e+008**
3	**1**	**22**	**2.01228e+008**
4	**1**	**6**	**2.00596e+008**
5	**2**	**5**	**2.00446e+008**
6	**2**	**1**	**2.00394e+008**

6 iterations, total sum of distances = 2.00394e+008 (Pixel or Bit)

في الجدول المرقم (1) يمثل العمود الأول التكرار لخوارزمية العنقدة أما العمود الثاني فيه الطور الذي على أساسه يتم تجزأة البيانات التابعة للصورة المدخلة. العمود الثالث يمثل عدد العناقيد الناتجة في كل تكرار أما العمود الأخير فيعني مجموع المسافات بين العناقيد لكل تكرار. كذلك بالنسبة للجدولين (2) و(3).



الشكل (2). يمثل الصورة الثانية التي طبقت عليها الخوارزمية اللامعلمية لتقدير دالة صورة إيماءة اليد الاعتيادية مع تقنية العنقدة بمعدل k.

جدول (2). يمثل مراحل تطبيق أسلوب العنقدة بمعدل k (k-mean) لصورة إيماءة اليد الاعتيادية في الشكل (2)

iter	phase	num	sum
1	1	100	1.1255e+007
2	1	13	9.6104e+006
3	1	5	9.23469e+006
4	1	1	9.21424e+006
5	2	0	9.21424e+006

5 iterations, total sum of distances = 9.21424e+006 (Pixel or Bit)

6. خوارزمية تجزئة وتقدير دالة كثافة الاحتمال لصورة إيماءة اليد المشوهة

Segmentation and Estimator Function Algorithm for Warping Hand Gesture Image

فيما يأتي خطوات الخوارزمية المقترحة:

الخطوة الأولى: قراءة الصورة المشوهة الرمادية بعد أخذها من مقطع فيديوي وتطبيق خوارزمية Beier-Neely للتشويه عليها وكالآتي:

خوارزمية Beier-Neely:

إن خوارزمية (Beier-Neely) تعتمد على استخدام الأسلوب العكسي من أجل إجراء عملية التشويه المستندة على التحديد المسبق للميزة إذ تعتبر هذه الخوارزمية من الخوارزميات الشعبية ذات المجال الواسع فقد قدمت هذه الخوارزمية من قبل Beier و Neely ونشرت في عام (1992) (والتي تعتبر الميزة أساس التحويل في الصور) [12]. لذا يمكن أن نعرف خوارزمية Beier-Neely بوصفها تقنية تستخدم من أجل عمل التشويه

والتركيب للصور ثنائية الأبعاد لهذا السبب تدعى باسم حقل تركيب الصور وأن أهم الخطوات الرئيسة لهذه الخوارزمية هو العمل على تحديد الميزات للصور المدخلة وتكون على نوعين:

- خوارزمية Beier-Neely لمتطابقة الخط الوحيد (في حالة الخط الواحد).

Single line correspondence Beier-Neely algorithm.

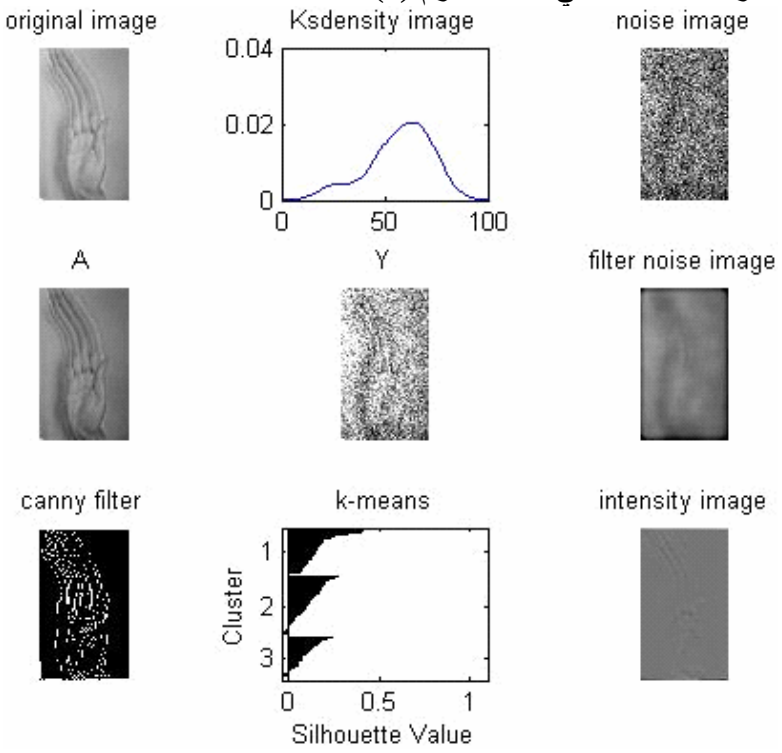
- خوارزمية Beier-Neely لمتطابقة الخطوط المتعددة (في حالة الخطوط المتعددة).

Multiple line correspondence Beier-Neely algorithm

إن خوارزمية حقل التشويه (في حالة الخط الواحد) يجب إن تحدد فيها بعض الأشياء الضرورية بالنسبة للصورة المدخلة وأيضا للصورة الناتجة.

الخطوات من الثانية إلى الثالثة عشر مشابهة تماما لنظيراتها في الخوارزمية السابقة:

وتطبيق الخوارزمية المقترحة أعلاه ممثل في الشكل المرقم (3).



الشكل (3). يمثل الصورة المشوهة التي طبقت عليها الخوارزمية اللامعلمية لتقدير دالة كثافة الاحتمال لصورة إيماءة اليد المشوهة مع تقنية العنقدة بمعدل k.

جدول (3). يمثل مراحل تطبيق أسلوب العنقدة بمعدل k (k-mean) لصورة إيماءة اليد المشوهة في الشكل(3)

iter	phase	num	sum
1	1	100	1.20482e+007
2	1	6	1.09183e+007
3	1	6	9.66603e+006
4	1	3	9.24697e+006
5	1	1	9.10073e+006
6	1	1	9.10073e+006
7	1	1	9.03223e+006
8	2	1	9.01808e+006

8 iterations, total sum of distances = 9.01808e+006 (Pixel or Bit)

7. تفسير النتائج:

إن أهم جزء من النتائج التي ظهرت لدينا هو أنه عند استخدام الدالة اللبية (Kernel) على الصورة المدخلة فإن المقدر اللبي قام بتقدير دالة الكثافة الاحتمالية لبيانات صورة إيماءة اليد، وبذلك أظهرت النتائج أن الدالة التي تخضع لها صورة الإيماءة مقاربة لدالة التوزيع الطبيعي وهذا يعني أنه تم تمييز الصورة على أنها حاوية على إيماءة اليد كما في الشكل (1) الذي يبين أن المنحنى قريب من منحنى التوزيع الطبيعي، أما في الشكل (2) و(3) فإن المنحنى يطابق منحنى التوزيع الطبيعي وهذا يعني أنه تم تمييز الصورة على أنها حاوية على صورة إيماءة اليد .

وكذلك تم تمييز الصورة المدخلة على أنها صورة حاوية على إيماءة اليد من خلال استخدام تقنية (K-Means) وذلك حسب الشكل الظاهر من خلال الرسم، إذ تم إيضاح أن الرسم ظهر بثلاث مناطق واضحة كعناقيد حسب التقسيمات المدخلة (الأصابع وراحة اليد والفراغات الموجودة في اليد)، وأظهرت النتائج في الشكل (2) والشكل (3) المناطق الثلاث، أما في الشكل (1) فإن الرسم الظاهر يظهر قيم شاذة لوجود خيال للإيماءة الموجودة في الصورة المدخلة، وكذلك نلاحظ وجود الخيال من خلال استخدام المحدد (Canny) إذ تم تحديد الخيال مع الإيماءة المدخلة.

8. الاستنتاجات Conclusions :

1. نلاحظ عند استخدام المقدر اللبي Kernel على صورة إيماءة اليد أن المقدر قام بتنعيم بيانات الصورة وتسويتها، وظهر منحني البيانات مشابهاً لمنحني التوزيع الطبيعي، وبذلك تبين لنا أن مقدر الكثافة اللامعلمية (المقدر اللبيKernel Estimator) لصورة إيماءة اليد قام بتمييز الصورة بأنها حاوية على إيماءة اليد.
2. يمكن استخدام تقنية (K-Means) لتمييز إيماءة اليد من خلال إدخال تقسيمات إيماءة اليد إلى دالة (K-Means).

نوصي باستخدام الشبكات العصبية مع نماذج ماركوف المخفية في التعرف على إيماءة اليد.

المصادر

[1] حمو، مناف يوسف، (2000)، "مقارنة مقدرات Kernel اللامعلمية لتقدير دوال الانحدار"، رسالة ماجستير غير منشورة، قسم الإحصاء والمعلوماتية، كلية الإدارة والاقتصاد، جامعة بغداد، العراق.

[2] ذنون، باسل يونس، (1997)، "التقدير أللبي" أسلوب بياني في التقدير الإحصائي"، بحث منشور علوم الرافدين، المجلد 9 ، العدد 1 ، ص 89-99.

[3] قاسم، عمر صابر، (2005)، "استخدام شبكة الانحدار المتدرج المحسنة في تشخيص مرض حصى الكلى"، رسالة ماجستير غير منشورة، كلية علوم الحاسوب والرياضيات، جامعة الموصل، العراق.

[4] Gasser, T., Muller, H.G. & Mammitzsch, R. (1985) ” Kernels for Nonparametric curve Estimation ” J.Royal S.S., series B, Vol.47, No.2, PP238-252.

[5] Hardle, W., (1990): "Applied nonparametric regression", Cambridge, Ma: Cambridge University press.

[6] MATLAB 6.5 The Language of Techical Computing, (2002): "Matlab Help", Math Works Inc.

[7] Parzen, E. (1962) , ”On estimation of a probability density function and mode“ Ann. Stat. Vol.33, PP1065-1076.

[8] Rabiner, L. R. (1989). "ATutorial on Hidden Markov Models and Selected Applications in Speech Recognition " , Proceeding of IEEE, Vol. 77, No. 2,PP. 257-286.

[9] Schucany, W.R. & Sommers, J. (1977) ” Improved of kernel type density estimators “ JASA, Vol.72, No.353, PP.420-423.

[10] Wilson, A.D. and Bobic, A.F.,(1997), "A state Based Approach to the Representation and Recognition of Gesture", Proceedings of IEEE, Vol.19, No.12, PP.1325-1337.

[12] Wolberg, G. (1998), " Image Morphing a Survey ", Dept. of Computer Science New York, NY 10031, No.14. pp 360-375, USA.