جامعة السودان للعلوم والتكنولوجيا

كلية الدراسات العليا

التَوُّف على أنماط سلاسل الأرقام الثنائية الناتجة عن ملفات الصورة والصوت والنص.

إعداد الطالب

سيد جعفر عبدالله الحسن

بكالوريوس علوم الحاسوب والإحصاء

إشراف

البروفيسور : عز الدين محمد عثمان

قدمت هذه الرسالة لنيل درجة الماجستير في علوم الحاسوب

بكلية الدراسات العليا في جامعة السودان للعلوم والتكنولوجيا

يونيو ٢٠٠٥ م

شكر وتقدير

قبل كل شي أود أن أبعث أسمى آيات الشكر والتقدير للبروفيسور عز الدين محمد عثمان ، الذي اقترح موضوع البحث . ثم أشكره مجدداً على صبره على كل المقترحات وكل التقصير الذي حصل مني طيلة هذه الفترة . كذلك أشكره على توجيهاته ونصائحه القيمة ، وأتمني من الله القدير أن يجعله ذخراً لهذا السودان، وأن يجعل علمه نافعاً .

ثم يطيب لي أن اشكر الأستاذ أحمد بابكر حسن ، بمكتبة معهد الإدارة العامة بالرياض (السعودية) ، الذي أتاح ، وهيئ لي كثيرٍ من المصادر العلمية ، من كتبٍ ، ومراجعٍ ، و إنترنت ودوريات .

كما أخص بالشكر مدير جامعة السودان للعلوم والتكنولوجيا ، وعميد كلية الدراسات العليا، و مسجل كلية الدراسات العليا، لإتاحة الفرصة لي ، لكي أكون طالباً في هذه الجامعة العملاقة .

كذلك أشكر مسبقاً ، كل من يطلع على هذا الجهد المتواضع ، وأشكر كل من يساهم في ترقية وتطوير هذا النوع من النظم(نظم التَّعرُّف على الأنماط) ، والتي أعتقد أنها مفيدة إذ أنها مثار اهتمام كل الدول المتقدمة في هذا العصر.

ملخص البحث

يهدف البحث إلي تحديد السمات والملامح المميزة، لأنماط سلاسل الأرقام الثنائية للصورة وللصوت وللنص ، وبهذه السمات نستطيع توصيف تلك الأنماط من السلاسل . ويهدف البحث كذلك إلي تصميم نموذج، يمكننا من التَّعرُّف على هوية هذه الأنماط، ومن ثم تصنيفها إلي صنف الصورة أو صنف الصوت أو صنف النص .

في هذا البحث تمت دراسة أساليب وتقنيات معالجة الأنماط ، واستعملت هذه الأساليب لتحليل أنماط من سلاسل الأرقام الثنائية للصور وللصوت وللنص . واستخدمت عدد من أساليب تحليل النسيج، لاستخلاص سمات وملامح تلك الأنماط . ثم استخدم نموذج (*Linear Discriminant Analysis*) لانتقاء أفضل السمات المميزة بين الأصناف الثلاثة : صنف سلاسل الأرقام الثنائية للصور، وصنف سلاسل الأرقام الثنائية للصوت، وصنف سلاسل الأرقام الثنائية للنصوص، و تم تطبيق معيار (*Mahalanobis distance*) لقياس الاختلافات بين هذه الأصناف، و تمت عملية البحث عن أهم، وأفضل السمات المميزة باستعمال طريقة البحث المتدرج(*stepwise*). ومن هذه السمات المميزة التي تم اختيارها، تم تكوين دوال للتمييز بين الأصناف، بناءً على معيار *Fisher*، واختبرت هذه الدوال، لقياس مدي قدرتها على التصنيف، والتعرف على الأنماط .

من أهم النتائج التي أنجزت في هذا البحث، أن هنالك فرق واضح، ومعتبر بين أصناف سلاسل الأرقام الثنائية للصورة وللصوت وللنص . و كذلك تم التوصل إلى إمكانية تصنيف نسبة ٨١.٣% من عينات التدريب، و التعرف على نسبة ٧٣.٧% من عينات الاختبار .

Abstract

The main objective of this thesis is to extract and select the discriminatory features of a stream of bit patterns for picture, sound, and text. These features should describe the stream of these patterns, and ultimately provide a pattern recognition system that identifies and classifies the new patterns into their relevant classes

Some techniques of **pattern recognition** methods, and **texture** analysis methods have been used and tested to extract the pattern features. **Linear Discriminant Analysis** method has been used to select the relevant discriminatory features according to **Mahalanobis distance** measure among pattern classes. A searching procedure has been performed using the **stepwise search** method. Finally the discriminant functions based on **Fisher's** criterion have been created to identify the class pattern.

Results have shown significant differences in a stream of bit patterns of picture, sound, ant text. About 81.3% of the **original groups** cases have been correctly classified, and 73.7% of the **test groups** cases have been correctly recognized.