ملخص الرسالة

إن معالجة الإشارات الرقمية (دي إس بي) مميزة عن باقي التطبيقات الأخرى في عِلْمِ الحاسبات من حيث النوعِ الفريدِ للبياناتِ المستخدمة (الإشارات). في أغلب الحالات، هذه الإشارات تَنشأُ كبيانات حسية من العالم الحقيقي مثل الاهتزازات الزلزالية، صور بصرية، موجات صوتية، الخ. معالجة الإشارات الرقمية هي الرياضيات، الخوارزميات، والتقنيات المستخدمة لمعالجة هذه الإشارات بعد تحويلها إلى الشكل الرقمي. من خلال التطور التاريخي لاستعمال الحاسبات، تطبيقات معالجة الإشارات الرقمي ة دفعت إلى حدود قوة الحساب ، خصوصاً من ناحية حساب الوقت الحقيقي. بينما تَراوحت الإشارات المستعملةَ بشكل واسع من إدارة الوسائط بالصوت ، تسجيل صوتي، وفيديو إلى الرادارِ المتَخصصِ وبيانات السونارِ، معظم الحسابات المنفذة بأنظمة معالجة الإشارات لها نفس الخصائصِ الحسابية الأساسية. طبيعة تَوازي البيانات الموجود في العديد من وظائف معالجة الإشارات الرقمية جعلت خوارزميات معالجة الإشارات مثالية لتمثيلها عن طريق هاردوير.

بينما التطبيقات شملت حيزا واسعا، فان البارامترات الحسابية الأساسية لمعظم عمليات معالجة الإشارات الرقمية تبقى بدون تغيير: الحاجة لأداءِ في زمن قياسي ضمن البارامترات المعطاة لهذا النظام، وفي معظم الحالات نحتاج إلى تعديل البيانات وشروط الحساب. في هذا السياقِ، مصفوفة بوابات المجال القابلة للبرمجة (إف بي جي أي) تكون انسب في مثل هذه الحالات . حيث هذه الحاسبات المعتمدة علي برمجة الخانة توفر كميات كافية مِن البوابات والمسجلات التي يمكن من السهولة تعديلها لدعم التَوازي في العديد من تطبيقات معالجة الإشارات الرقمية وخصوصا السعة الحالية لهذه الآلات من البوابات تَتجاوز مليون بوابة لكل أداةِ.

في هذه الأطروحة تمت دراسة اثنان من تطبيقات معالجة الإشارات الرقمية المعتمدة على معالجة الصوت. هذان التطبيقان هما التعرف على المتكلم وتمييز الصوت. حيث أن هذان التطبيقان لهما تطبيقات كثيرة في الحياة اليومية مثل النظم الآمنة ،التسوق من بعد، وإدارةِ المعرفة. ولكن هذان التطبيقان (التعرف على المتكلّمَ وتمييز الصوتَ) يستغرقان وقتا طويلا في حالة استعمال المعالجِ التقليديِ وهم في حاجة لاستخدام الهاردوير للتطبيقات الحية ذات وقت قصير جدا.

تمييز الصوت آليا (أي إس آر) هو العملية التي بواسطتها بتم تحويل إشارة سمعية إلى النَصِ. على وجه التحديد، يمكِن أَن يعتبر تمييز الصوت آليا هو مشكلة تَحويل خطاب صوتي إلى الكلمات مقروءة. حيث أن إشارةِ هذا الخطاب تحول إلى الوحدات الأولية من الخطاب مثل الفونيمات أَو الكلمات وتفسيرِ السلسلة المحولة لكي يسمحا لتصحيح الكلمات المعترف بها لفهم الخطاب. مهام تمييز الصوت آليا يمكن أَن تصنف إلى التابعِ المتوافقِ والأصناف المستقلة. في الأنظمة المعتمدة على المتكلم، المحقق يكون مصمم لمتكلم وحيد. بينما المحقق في حالة المتكلمِ المستقلِ يتم تدريب النظام من قبل عديد من المتكلمينِ و يستخدم أيضا من قبل العديد من المتكلمين. ويعتبر نظام ماركوف الغير مرئي(إتش إم إم) هو إحدى أكثر الخوارزميات الناجحة المستخدمة في تمييز الصوت. نظام ماركوف الغير مرئي يعتمد على تقنية عرض احتمالية التي تَستعمل عدد محدود من الحالات والانتقالات الرسمية

المرتبطة بها ليحاكي تغيرات الإشارات. تم استعمال هذا النظام على نطاق واسع لتَشكيل وحدات الصوت الأساسية في تمييز الصوت.

قمنا بتعديل الخوارزميةَ الأماميةَ، التي هي قلب نظام ماركوف ، لتمثيلها علي هاردوير . قمنا بتصميم الخوارزم المعدل باستخدام لغة (في إتش دي إل). وتم تطبيقه باستخدام مصفوفة البوابات القابلة للبرمجة(إف بي جي) وإختبرنَا تطبيقَنا باستعمال أحد كروت شركة Xilinx (STARTER Kit). تصميمنا يشغل 35 % (في حالة قيم العدد الصحيحِ) و 81 % (في حالة النقطة العائمة) من مصادر الشريحة الموجودة على الكارت (Spartan3XC3S200-4). حققنا دقة تعرف بنسبة 98 % في زمن قدره 1.62 مايكرو ثانية.

إن التعرف على المتحدث آليا هو أحد حقول تمييز المتكلم الذي يعتبر مشكلة مجاراة النمط . أنظمة التعرف على المتكلمِ آليا تنقسم إلى قسمين رئيسيينِ: معتمد على نَص ونَص مستقل. في الأنظمة المعتمدة على النَص، النصوص المستخدمة للتدريب والاختبار واحدة . من الناحية الأخرى في أنظمة النَص المستقل، النص المستعمل في مرحلة الاختبار ليس بالضرورة هو نفسه المستعمل في مرحلةِ التدريب. إن الشبكات العصبيةَ ثبتت ناجح ا في مرحلة مجاراة النمط بسبب قابليةِ تعرفها الفائقة على الأنماط . القابلية العالية للشبك ات العصبية في التصنيف الدقيق جعلها مناسبة لحل مشاكلِ تصنيف النمط المعقدة للصوت . علاوة على ذلك، طبيعة التَوازي للعمليات للشبكات العصبية يمكن أَن تسرع حساب أعداد كبيرة من مجاراة النمط.

بَنينَا الهاردوير لإحدى الشبكات العصبية للتعرف على المتكلم باستخدام لغة (في إتش دي إل). عد لنَا دالة التحجيم(دالة غير خطية) . قمنا ببناء واختبار تطبيقَنا باستخدام أحد كروت شركة Xilinx. بسبب المصادرِ المحدودة للكارت (Starter Kit) إستعملنَا صيغةَ النقطة العائمة المخفضة لتَمثيل الأعداد، خلايا عصبية إفتراضية، وتقنية معايير متعددة الإرسال. تصميمنا يشغل حيزا 76 % من مصدرِ الشريحة (Spartan-3XC3S200-4 الموجودة على الكارت . إن البناء للشبكة العصبيةِ نموذج ، ويمكن بسهولة زيادة أَو نقص عددَ الخلايا العصبية وكذلك الطبقاتِ.

تصميمنا قدم ليقوم بالتعرف على المتكلّم باستخدام الخلايا الشبكية متبوعا بالتعرف على الكلام باستخدام نظام ماركوف الغير مرأى . أثبتت مصفوفة بوابات المجال القابلة للبرمجة كفاءة في تحسين أداءِ كلتا التطبيقين(متكلّم وتمييز الصوت).