الجمهورية الشعبية الديمقراطية الجزائرية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
المدرسة العليا للإعلام الآلي 08 ماي 1945. بسيدي بلعباس
École Supérieure en Informatique
-08 Mai 1945- Sidi Bel Abbès

THESIS

To obtain the diploma of **Master**
Field: **Computer Science**
Specialty: **Ingénierie des Systèmes Informatiques (ISI)**

Theme

Exploring the Current State of Intelligent Multimodal Human-Computer Interaction Systems

---

Presented by:
Tebabna Ahmed Rami

Submission Date: **XX, 2024**
In front of the jury composed of:

Mr. Someone	President
Ms. Bedjaoui Mohamed	Supervisor
Mr. Someone	Examiner

Academic Year : 20023/2024

Abstract

This thesis explores the current state of intelligent multimodal Human-Computer Interaction (HCI) systems, emphasizing gaze estimation, gesture recognition, and speech recognition. The problem addressed is the need to understand the advancements, challenges, and opportunities in integrating these technologies to enhance user interaction. The objective was to conduct a comprehensive literature review to identify recent progress, key issues, and potential applications. The findings highlight significant improvements in computer vision algorithms, real-time processing capabilities, and the integration of machine learning models, improving accuracy and responsiveness. In conclusion, this study provides valuable insights and a foundation for future research, underscoring the potential of multimodal HCI systems to create more intuitive and efficient user interactions.

Keywords— Human-computer interaction, distributed systems, deep learning, Natural language processing, gestures recognition, voice user interface

الملخص

يستكشف هذا البحث الحالة الحالية لأنظمة التفاعل بين الإنسان والحاسوب (HCI) متعددة الوسائط الذكية، مع التركيز على تقدير النظرة، التعرف على الإيماءات، والتعرف على الكلام. تتمثل المشكلة التي تم معالجتها في الحاجة إلى فهم التقدمات والتحديات والفرص في دمج هذه التقنيات لتعزيز التفاعل مع المستخدم. كان الهدف هو إجراء مراجعة أدبية لتحديد التقدم الحديث والقضايا الرئيسية والتطبيقات المحتملة. تبرز النتائج تحسينات كبيرة في خوارزميات الرؤية الحاسوبية، وقدرات المعالجة في الوقت الفعلي، ودمج نماذج التعلم الآلي، مما يحسن الدقة والاستجابة. في الختام، يوفر هذا البحث رؤى قيمة وأساسًا للبحث المستقبلي، مشددًا على إمكانات أنظمة HCI متعددة الوسائط في خلق تفاعلات مستخدم أكثر بديهية وكفاءة.

الكلمات المفتاحية: التفاعل بين الإنسان والحاسوب، الأنظمة الموزعة، التعلم العميق، معالجة اللغة الطبيعية، التعرف على الإيماءات، واجهة المستخدم الصوتية.

Résumé

Cet thèse explore l'état actuel des systèmes intelligents d'interaction homme-machine (HCI) multimodaux, en mettant l'accent sur l'estimation du regard, la reconnaissance des gestes et la reconnaissance vocale. Le problème abordé est la nécessité de comprendre les avancées, les défis et les opportunités dans l'intégration de ces technologies pour améliorer l'interaction utilisateur. L'objectif était de mener une revue de littérature complète pour identifier les progrès récents, les principaux enjeux et les applications potentielles. Les résultats mettent en évidence des améliorations significatives des algorithmes de vision par ordinateur, des capacités de traitement en temps réel et de l'intégration de modèles d'apprentissage automatique, améliorant la précision et la réactivité. En conclusion, cette étude fournit des perspectives précieuses et une base pour les recherches futures, soulignant le potentiel des systèmes HCI multimodaux pour créer des interactions utilisateur plus intuitives et efficaces.

Mots clés: Interaction homme-machine, systèmes distribués, apprentissage profond, traitement du langage naturel, reconnaissance des gestes, interface utilisateur vocale