

Diss. ETH No. 13849

Orientable Edge Touch Screen: Implementation and Software for Teaching Geometry to the Blind

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH
(ETH Zürich)

for the degree of
Doctor of Natural Sciences

presented by

Pierre-Alain Genilloud

Dipl. Phys. ETH

born March 14, 1971

citizen of Villarepos, FR

Prof. Dr. Klaus Hepp, examiner
Prof. Dr. Rodney Douglas, co-examiner

2000

Abstract

We consider in this thesis the representation of graphical data and images to the touch sense. This encompasses the conception and implementation of a touch screen, a mechanical device able to produce a relief representation fitting the touch characteristics, and the development of the software to extract the relevant information from graphical data or images, in order to construct from them a touch adapted representation.

We discuss the potential of touch for the recognition of graphical representations. We consider the different modes of data acquisition which touch can adopt, and point out the advantage of parallel over sequential acquisition.

These observations, as well as tests of different candidates for a touch representation, led us to look for reliefs permitting a full “hand grasp”. These showed the best performances, allowing a recognition by touch of relatively complex figures.

After having considered different possible implementations, we adopt the Orientable Edge Touch Screen (OETS), which provides an efficient representation of images and graphics to the touch sense, while allowing to avoid the technical problems imposed by the other candidates.

We present for this screen our hardware, electronics and software solution for its implementation and control.

The next theme is the computer processing of images or graphics, allowing the extraction and projection of the relevant information on the touch screen.

In a chapter, we treat relatively simple images, grabbed from a cartoon, which allows us to determine the level of image understanding a computer system has to reach, in order to produce adequate touch representations. A clean extraction of the main objects silhouettes seems necessary. This kind of task is an unsolved problem of computer vision, for a unrestricted set of images.

We present a method based on Maximum Entropy, allowing the fusion of several cues for image segmentation and the extraction of the main contours. This method allows to get relatively good results, but they are still not sufficient for projection on our touch screen.

We consider also the more synthetic problem of representing geometrical constructions, bound to different euclidean geometry theorems. We define an error functional measuring the discrepancy between the ideal construction and its instantiation on the OETS. We use the degrees of freedom a geometrical construction provides, in order to reach lower error costs. This implies an information compression with loss, where the relevant information is preserved. We consider also the problem of determining these degrees of freedom automatically, given a single finite precision figure. This problem corresponds to the inference of the geometrical construction class from which the given instance was generated. This can be interpreted as an extraction of relevant features or as an object class recognition.

Our solution is efficient, making a good use of the resolution at which the base figure is given. It allows also to define the most important non-generic configurations, which have to be avoided when presenting a typical instance of a construction class to a student.

This approach is derived from an inference method called Minimal Minimum Description Length (MMDL), based on algorithmic complexity theory, which we have developed and which we present in appendix. We show there also its relationship with the Minimum Description Length principle (MDL) and prove some convergence properties as well as a link with the Maximum Entropy principle.

Résumé

Nous considérons dans cette thèse la représentation de données graphiques et d'images pour le toucher. Cela comprend la conception et l'implémentation d'un écran tactile, un appareil mécanique pouvant générer une représentation en relief convenant aux caractéristiques du toucher, et le développement du logiciel capable d'extraire l'information essentielle de données graphiques ou d'images, de manière à en construire une représentation tactile adaptée.

Nous discutons du potentiel du toucher pour la reconnaissance de représentations graphiques. Nous considérons les différents modes d'acquisition que le toucher peut adopter et soulignons l'avantage de l'acquisition parallèle par rapport à l'approche séquentielle. Ces observations, aussi bien que des tests sur différents candidats de représentations, nous ont amené à rechercher des reliefs permettant une prise complète de la main. Ceux-ci ont montré les meilleures performances, permettant la reconnaissance tactile de figures relativement complexes.

Après avoir considéré différentes implémentations possibles, nous adoptons un Écran Tactile à Contours Orientables, qui présente une représentation efficace d'images et de graphiques, tout en permettant d'éviter les problèmes techniques imposés par les autres candidats.

Nous présentons pour cet écran notre solution hardware, électronique et software, pour son implémentation et son contrôle.

Le sujet suivant est le traitement par ordinateur d'images ou de graphiques, permettant l'extraction et la projection de l'information essentielle sur l'écran tactile.

Dans un chapitre, nous traitons des images relativement simples, tirées d'un dessin animé, ce qui nous permet de déterminer le niveau de compréhension d'image que doit atteindre un ordinateur, pour générer des représentations tactiles adéquates. Une extraction claire des silhouettes des objets principaux semble nécessaire. Ce genre tâche est un problème non résolu de la vision par ordinateur ("computer vision"), pour un ensemble d'images non restreint.

Nous présentons une méthode basée sur le principe d'Entropie Maximale, permettant la combinaison de différents types d'information, pour la segmentation de l'image et l'extraction des contours principaux. Cette méthode permet d'obtenir de relativement bons résultats, mais ils restent insuffisants pour une projection sur notre type d'écran.

Nous considérons également le problème plus synthétique de la représentation sur écran tactile de constructions géométriques, liées à différents théorèmes de la géométrie euclidienne. Nous élaborons une fonctionnelle définissant un coût pour la divergence entre la construction idéale et sa réalisation sur l'Écran Tactile à Contours Orientables. Nous mettons en jeu les degrés de liberté inhérents aux constructions géométriques, de manière à atteindre des coûts de divergence plus bas. Ceci implique une compression d'information avec perte, où l'information essentielle est préservée. Nous considérons également le problème de déterminer ces degrés de liberté de manière automatique, sur la base d'une seule figure, définie avec une précision finie. Ce problème correspond à inférer la classe de constructions géométriques de laquelle la figure présentée a été tirée. Ceci peut être interprété comme une extraction des caractéristiques essentielles ou une reconnaissance de classe d'objet.

Notre solution est efficace, tirant un bon parti de la résolution avec laquelle la figure de base est proposée. Elle permet également de définir les configurations non génériques les plus importantes, qui doivent être évitées lorsque l'on veut présenter un exemple caractéristique d'une classe de figures à un élève.

Cette approche est dérivée d'une méthode d'inférence appelée "Minimal Minimum Description Length principle" (MMDL), basée sur la théorie de la complexité algorithmique, que nous avons développée et que nous présentons en appendice. Nous y montrons également sa parenté avec le principe de longueur de description minimale (MDL) et démontrons quelques propriétés de convergence ainsi qu'un lien avec le principe d'Entropie Maximale (ME).