

UNIVERSITÀ DEGLI STUDI DI SALERNO
DIPARTIMENTO DI INFORMATICA

DOTTORATO DI RICERCA IN INFORMATICA
XII CICLO - NUOVA SERIE



Tesi di Dottorato in

New Methods, Techniques and Applications for Sketch Recognition

Mattia De Rosa

Ph.D. Program Chair
Prof. Giuseppe Persiano

Supervisors
Prof. Gennaro Costagliola

Dott. Vittorio Fucella

ANNO ACCADEMICO 2013/2014

Abstract

The use of diagrams is common in various disciplines. Typical examples include maps, line graphs, bar charts, engineering blueprints, architects' sketches, hand drawn schematics, etc.. In general, diagrams can be created either by using pen and paper, or by using specific computer programs. These programs provide functions to facilitate the creation of the diagram, such as copy-and-paste, but the classic WIMP interfaces they use are unnatural when compared to pen and paper. Indeed, it is not rare that a designer prefers to use pen and paper at the beginning of the design, and then transfer the diagram to the computer later.

To avoid this double step, a solution is to allow users to sketch directly on the computer. This requires both specific hardware and sketch recognition based software. As regards hardware, many pen/touch based devices such as tablets, smartphones, interactive boards and tables, etc. are available today, also at reasonable costs. Sketch recognition is needed when the sketch must be processed and not considered as a simple image and it is crucial to the success of this new modality of interaction. It is a difficult problem due to the inherent imprecision and ambiguity of a freehand drawing and to the many domains of applications. The aim of this thesis is to propose new methods and applications regarding the sketch recognition. The presentation of the results is divided into several contributions, facing problems such as corner detection, sketched symbol recognition and autocompletion, graphical context detection, sketched Euler diagram interpretation.

The first contribution regards the problem of detecting the corners present in a stroke. Corner detection is often performed during preprocessing to segment a stroke in single simple geometric primitives such as lines or curves.

The corner recognizer proposed in this thesis, RankFrag, is inspired by the method proposed by Ouyang and Davis in 2011 and improves the accuracy percentages compared to other methods recently proposed in the literature.

The second contribution is a new method to recognize multi-stroke hand drawn symbols, which is invariant with respect to scaling and supports symbol recognition independently from the number and order of strokes. The method is an adaptation of the algorithm proposed by Belongie et al. in 2002 to the case of sketched images. This is achieved by using stroke related information. The method has been evaluated on a set of more than 100 symbols from the *Military Course of Action* domain and the results show that the new recognizer outperforms the original one.

The third contribution is a new method for recognizing multi-stroke *partially* hand drawn symbols which is invariant with respect to scale, and supports symbol recognition independently from the number and order of strokes. The recognition technique is based on subgraph isomorphism and exploits a novel spatial descriptor, based on polar histograms, to represent relations between two stroke primitives. The tests show that the approach gives a satisfactory recognition rate with partially drawn symbols, also with a very low level of drawing completion, and outperforms the existing approaches proposed in the literature. Furthermore, as an application, a system presenting a user interface to draw symbols and implementing the proposed autocompletion approach has been developed. Moreover a user study aimed at evaluating the human performance in hand drawn symbol autocompletion has been presented. Using the set of symbols from the *Military Course of Action* domain, the user study evaluates the conditions under which the users are willing to exploit the autocompletion functionality and those under which they can use it efficiently. The results show that the autocompletion functionality can be used in a profitable way, with a drawing time saving of about 18%.

The fourth contribution regards the detection of the graphical context of hand drawn symbols, and in particular, the development of an approach for identifying *attachment areas* on sketched symbols. In the field of syntactic recognition of hand drawn visual languages, the recognition of the relations

among graphical symbols is one of the first important tasks to be accomplished and is usually reduced to recognize the attachment areas of each symbol and the relations among them. The approach is independent from the method used to recognize symbols and assumes that the symbol has already been recognized. The approach is evaluated through a user study aimed at comparing the attachment areas detected by the system to those devised by the users. The results show that the system can identify attachment areas with a reasonable accuracy.

The last contribution is EulerSketch, an interactive system for the sketching and interpretation of Euler diagrams (EDs). The interpretation of a hand drawn ED produces two types of text encodings of the ED topology called *static code* and *ordered Gauss paragraph* (OGP) code, and a further encoding of its regions. Given the topology of an ED expressed through static or OGP code, EulerSketch automatically generates a new topologically equivalent ED in its graphical representation.

Abstract

L'uso di diagrammi è comune in varie discipline. Esempi tipici includono mappe, grafici lineari, grafici a barre, progetti ingegneristici, schizzi architettonici, schemi disegnati a mano, ecc. In generale i diagrammi possono essere creati o utilizzando carta e penna o specifici software. Questi programmi forniscono funzioni per facilitare la creazione del diagramma, come il copia e incolla, ma le classiche interfacce WIMP che utilizzano sono poco naturali se confrontate a carta e penna. Infatti non è raro che un progettista preferisca usare carta e penna all'inizio della progettazione e trasferisca in seguito il diagramma al computer.

Per evitare questo doppio passaggio una soluzione è quella di disegnare sketch direttamente al computer. Questo richiede sia hardware specifico sia software basato sul riconoscimento di sketch. Per quanto riguarda l'hardware, molti dispositivi touch come tablet, smartphone, lavagne e tavoli interattivi, ecc. sono oggi disponibili, anche a costi ragionevoli. Il riconoscimento di sketch è necessario quando lo sketch deve essere elaborato e non deve essere considerato solo una semplice immagine ed è cruciale per il successo di questa nuova modalità di interazione. Il riconoscimento è un problema difficile a causa dell'intrinseca imprecisione ed ambiguità del disegno a mano libera e dei numerosi domini di applicazione. Lo scopo di questa tesi è di proporre nuovi metodi ed applicazioni riguardo il riconoscimento di sketch. La presentazione dei risultati è divisa in vari contributi che affrontano problemi quali l'identificazione degli angoli, il riconoscimento e l'autocompletamento di simboli disegnati a mano libera, la graphical context detection, l'interpretazione di diagrammi di Eulero disegnati a mano libera.

Il primo contributo riguarda il problema di identificare gli angoli presenti

in uno stroke. L'identificazione degli angoli è spesso eseguita durante il preprocessing per segmentare uno stroke in singole semplici primitive geometriche come linee o curve. RankFrag, il riconoscitore di angoli proposto in questa tesi, è ispirato al metodo proposto da Ouyang e Davis nel 2011 e migliora le percentuali di accuratezza rispetto ad altri metodi recentemente proposti in letteratura.

Il secondo contributo è un nuovo metodo per riconoscere simboli multi-stroke disegnati a mano, che è invariante rispetto alla scala e supporta il riconoscimento di simboli indipendentemente dal numero e dall'ordine degli stroke. Il metodo è un adattamento dell'algoritmo proposto da Belongie et al. nel 2002 al caso degli sketch. Questo è stato ottenuto utilizzando informazioni relative allo stroke. Il metodo è stato valutato su un insieme di più di 100 simboli appartenenti al dominio *Military Course of Action* ed i risultati mostrano che il nuovo riconoscitore ottiene performance migliori di quello originale.

Il terzo contributo è un nuovo metodo per il riconoscimento di simboli multi-stroke *parziali* disegnati a mano, che è invariante rispetto alla scala e supporta il riconoscimento dei simboli indipendentemente dal numero e dall'ordine degli stroke. La tecnica di riconoscimento è basata su isomorfismo tra sottografi e utilizza un nuovo descrittore spaziale, basato su istogrammi polari, per rappresentare le relazioni tra due primitive (di uno o più stroke). I test mostrano che l'approccio offre tassi di riconoscimento soddisfacenti con simboli parzialmente disegnati anche con un livello molto basso di completamento del disegno ed ottiene performance migliori degli approcci esistenti proposti in letteratura. Inoltre, come applicazione, è stato sviluppato un sistema che presenta un'interfaccia utente per disegnare simboli e che implementa l'approccio di autocompletamento proposto. Inoltre è stato presentato uno studio avente lo scopo di valutare le performance umane nell'autocompletamento di simboli disegnati a mano. Utilizzando l'insieme di simboli appartenenti al dominio *Military Course of Action* lo studio valuta le condizioni sotto le quali gli utenti sono disposti a sfruttare la funzionalità di autocompletamento e quelle sotto le quali essi riescono ad utilizzarla in maniera efficiente. I risultati mostrano che la funzionalità di autocompletamento può essere usata

in maniera proficua con un risparmio del tempo di disegno di circa il 18%.

Il quarto contributo riguarda la graphical context detection di simboli disegnati a mano, ed in particolare, lo sviluppo di un approccio per identificare le *aree d'attacco* in simboli disegnati a mano libera. Nel campo del riconoscimento sintattico di linguaggi visuali disegnati a mano, il riconoscimento delle relazioni fra simboli grafici è uno dei primi task che deve essere compiuto ed è usualmente ridotto al riconoscimento delle aree d'attacco di ogni simbolo e delle relazioni tra loro. L'approccio è indipendente dal metodo utilizzato per riconoscere i simboli ed assume che il simbolo sia già stato riconosciuto. L'approccio è valutato mediante uno studio avente lo scopo di comparare le aree d'attacco individuate dal sistema con quelle indicate dagli utenti. I risultati mostrano che il sistema può identificare le aree d'attacco con una ragionevole accuratezza.

L'ultimo contributo è EulerSketch, un sistema interattivo per lo sketching e l'interpretazione dei diagrammi di Eulero (ED). L'interpretazione di ED disegnati a mano produce due tipi di codifica testuale della topologia di un ED chiamati *static code* ed *ordered Gauss paragraph (OGP) code* ed un'ulteriore codifica delle sue regioni. Data la topologia di un ED espressa mediante static code o OGP code, EulerSketch genera automaticamente un nuovo ED topologicamente equivalente nella sua rappresentazione grafica.