

# L'UTILIZZO DI METODI AUTOMATICI PER L'ANALISI DELLA FORMA APPLICATI ALLA RICERCA ARCHEOLOGICA

di Loris Nanni, Alessandra Lumini, Monica Toschi

Negli ultimi anni c'è stato un crescente interesse all'applicazione delle tecnologie informatiche al problema della conservazione e fruizione dei beni culturali, soprattutto in Italia, dove questo tema è particolarmente importante. In ambito archeologico l'applicazione di metodologie di analisi automatizzata a problemi di riconoscimento, classificazione o datazione dei reperti ha permesso nuove chiavi di lettura e risultati originali in problemi affrontati da molto tempo con metodologie classiche. Grazie a ciò è notevolmente cresciuto l'interesse degli studiosi di archeologia per l'informatica, vista ora non solo come scienza sussidiaria utile a migliorare la catalogazione e l'archiviazione dei reperti, ma come strumento in grado di coadiuvare la ricerca in tutti i suoi aspetti tramite potenti tecniche di intelligenza artificiale, computer vision e pattern recognition.

**N**egli ultimi anni l'attività di catalogazione e l'archiviazione in formato elettronico dei beni culturali ha richiesto grosso impegno e risorse, ma ha permesso nuovi metodi di fruizione (e.g. sistemi per la visita virtuale degli ambienti museali, visita ipermediale dei reperti), la nascita di nuove tecniche di restauro virtuale e l'uso di metodi automatici per l'analisi del patrimonio artistico e archeologico (Stanco, Tanasi, 2011). In ambito archeologico lo stretto legame che si è creato tra informatica e scienza archeologica ha preso il nome di Digital Archaeology e comprende tutti i possibili ambiti applicativi delle tecnologie informatiche all'archeologia.

Il problema, preso in esame in questo lavoro, riguarda la classificazione e la ricerca di reperti archeologici all'interno di grandi archivi digitali mediante tecniche di ricerca automatica basate sulla similarità di forma. I sistemi di ricerca automatici si basano sulla definizione di una rappresentazione dei reperti tramite l'uso di descrittori di forma, e l'uso di tecniche di apprendimento automatico opportunamente addestrate al riconoscimento e alla classificazione. L'importanza di un simile sistema non è limitato agli archeologi, ma può consentire agli studiosi d'arte, agli appassionati e ai curiosi, di fruire di un sistema che è in grado di ricercare e ordinare per similarità tutte le immagini di reperti più somiglianti a una fornita in ingresso come esempio.

Il problema della ricerca per similarità e della classificazione di reperti archeologici è tutt'altro che semplice e per questo motivo è stato tradizionalmente affrontato con tecniche semiautomatiche basate sull'annotazione testuale.

I grossi passi avanti fatti nell'ambito della visione artificiale hanno suggerito l'applicazione di tecniche basate sull'estrazione automatica di descrittori legati a caratteristiche percettive, quali i colori o la forma, e sull'uso di classificatori addestrati per problemi di ricerca e classificazione di

oggetti artistici e reperti archeologici (Lumini, 1999; Lumini, Maio, 1999; Krassimira 2011).

L'obiettivo dell'articolo sarà mostrare l'utilità di descrittori di forma in ambito archeologico. In particolare si mostrerà che l'algoritmo utilizzato, open source disponibile online <https://www.dei.unipd.it/node/2357>, può essere utilizzato senza modificare i suoi parametri in diverse applicazioni. Nella fattispecie testeremo il metodo in due problemi:

- classificazione dei motivi elementari delle ceramiche Kamares;
- classificazione di frammenti ceramici provenienti dal canale portuale della città di Classe.

Il dataset di ceramiche "Kamares" è stato creato nell'ambito del Progetto Archeomatica - Università di Catania (Stanco et al 2011; Guarnera et al 2011). Esso si riferisce alla "classe ceramica" di stile Kamares.

Mentre il secondo dataset contiene frammenti, principalmente anfore, provenienti dal canale portuale della città tarantoantica di Classe; in questo caso il sistema automatico deve fornire una corretta classificazione del reperto ad una delle diverse tipologie ([http://archaeologydataservice.ac.uk/archives/view/amphora\\_ahrb\\_2005/cat\\_amph.cfm](http://archaeologydataservice.ac.uk/archives/view/amphora_ahrb_2005/cat_amph.cfm)) relativa alla forma del corpo ceramico: orlo, collo, fondo, ecc..

Una descrizione più accurata dei dataset utilizzati è fornita in sezione nel prossimo paragrafo.

## I DATASET STUDIATI

### *La ceramica Kamares*

La ceramica Kamares è una tipologia di ceramica prodotta in una località dell'isola di Creta, la grotta sul Monte Ida, che si affermata il Minoico Medio I (2160 a. C.) e il Minoico Medio II (1800 a. C.) come tipo di decorazione vascolare.

I vasi venivano ottenuti dalla lavorazione al tornio girevole, questa lavorazione richiedeva un'alta preparazione a livello tecnico da parte del vasaio che riusciva a fare vasi molto sottili. La decorazione era caratterizzata da una policromia (bianco, rosso e arancione) su sfondo nero; il suo repertorio decorativo era ricchissimo: oltre alla spirale, decorazione tipica del repertorio cicladico, comparivano temi vegetali e marini, sempre comunque incentrati sull'elemento geometrico.

Nel Minoico Tardo I si affermò lo stile naturalistico caratterizzato da soggetti tratti da mondo naturale con una maggiore aderenza al soggetto reale, nonostante l'evidente semplificazione del soggetto. Qualche esempio dei motivi presenti nel dataset è mostrato in figura 1.



Fig. 1 - Esempi di motivi della ceramica Kamares.

#### Vasi del porto di Classe

Le anfore di cui parliamo in questo studio provengono dalle campagne di scavo effettuate nell'antico quartiere portuale di Classe (Boschi, 2011), in funzione già in epoca repubblicana ma che ebbe maggior importanza dal 27 a.C. quando Augusto decise di stanziarvi una flotta permanente.

A Classe non era praticata solo una navigazione marittima, ma anche una navigazione endolagunare che poteva essere effettuata grazie alla Fossa Augusta, che collegava Classe al Delta del Po.

Il porto di Classe fu protagonista di diversi rifacimenti, i più importanti sono: quello del 402, quando l'Imperatore romano Onorio decise di spostare la capitale dell'Impero Romano d'Occidente da Milano a Ravenna, e quello del 493, quando Teodorico depose Odoacre, in questa occasione il porto divenne un porto ad uso esclusivamente commerciale.

Il dataset analizzato in questo studio contiene anfore prodotte, per la maggior parte dei casi, nell'Africa del nord. Sono inoltre presenti alcuni casi provenienti dall'Italia meridionale e dalla Grecia (Bonifay, 2004). Nello scavo di Classe c'è stato anche il ritrovamento di un esemplare della tipologia Crypta Balbi di cui l'origine è sconosciuta, ma da studi sulla petrologia si può ipotizzare una provenienza italiana, greca o asiatica.

Le anfore sono databili dal II secolo d. C. al VII secolo d. C., con pochissimi casi di ritrovamenti di anfore precedenti a questo periodo. I prodotti trasportati erano olio di oliva, vino, prodotti impossibili da produrre a Classe per via della conformazione del terreno, e salsa di pesce (Keay, 1984).

Qualche esempio di oggetti (forma già estratta) appartenenti al dataset è mostrato in figura 2.

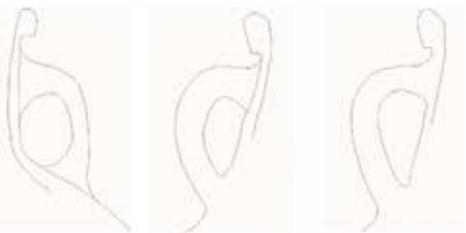


Fig. 2 - Esempi di cocci di anfore.

## UN SISTEMA DI CLASSIFICAZIONE AUTOMATICA BASATO SU DESCRITTORI DI FORMA

L'estrazione di descrittori di forma e la classificazione è stata eseguita mediante 6 passi di elaborazione illustrati in figura 3:

### 1. ESTRAZIONE DEI CONTORNI

Il primo passo del metodo riguarda l'estrazione dei contorni delle immagini che devono essere esaminate. Per rendere i descrittori di forma invarianti rispetto a rotazione e variazione di dimensione, le immagini da analizzare sono ritagliate e ridimensionate alle stesse dimensioni. I contorni sono estratti sia per l'immagine originale che per le sue versioni capovolte e ruotate (tutti gli angoli di rotazione con step di 22.5°). In questo modo si ottengono 32 diversi contorni dai quali estrarre i descrittori.

### 2. CALCOLO DI MATRICI DI RAPPRESENTAZIONE

Per ciascun contorno estratto al passo precedente vengono calcolate tre rappresentazioni matriciali: un istogramma bidimensionale denominato shape context (SC) e proposto da (Belongie et alii 2002); un metodo di rappresentazione di tale istogramma basato su calcolo delle distanze tra elementi strutturali, l'inner distance shape context method (ID), proposto da (Ling, Jacobs, 2007); un descrittore che misura le distanze dei punti di contorno dalle tangenti, denominato height function (HF) e illustrato in (Wang et alii 2012).

Alla fine di questo passo di elaborazione per ciascuna immagine di input sono state ottenute 3x32 matrici di rappresentazione (3 rappresentazioni x 32 contorni estratti nello step 2).

### 3. ESTRAZIONE DEI DESCRITTORI VETTORIALI

La rappresentazione matriciale non solo viene utilizzata direttamente per il confronto, ma costituisce anche un passo intermedio all'estrazione di ulteriori descrittori tessiturali. Ciascuna matrice viene considerata come se fosse un'immagine e descritta tramite un potente descrittore locale di tessitura: Local phase quantization (LQP), un operatore proposto da (Ojansivu, Heikkila 2008) e basato sulla proprietà di invarianza dello spettro di fase di Fourier.

L'estrazione dei descrittori LPQ non avviene sull'intera matrice ma localmente su ciascuna delle sotto-matrici ottenute dividendo la matrice in  $K1 \times K2$  regioni ( $K1=K2=4$  per ID e SC;  $K1=K2=2$  per HF; dato che HF è di dimensioni minori delle precedenti): per ciascuna regione vengono estratti 2 descrittori LPQ a diversi valori di raggio di analisi (3, 5), per dettagli su LPQ (Ojansivu, Heikkila 2008).

### 4. CONFRONTO

Il confronto tra due forme avviene calcolando la distanza standard per le matrici ID/SC/HF e utilizzando la distanza di Jeffrey per i descrittori LPQ (Nanni et alii, 2012). Per tutte e quattro le rappresentazioni vengono considerati tutti i diversi contorni ottenuti nello step 1.

### 5. NORMALIZZAZIONE E FUSIONE

Per ottenere il valore finale di classificazione i risultati ottenuti dal confronto dei diversi descrittori matriciali e di tessitura sono prima normalizzati a media 0 e deviazione standard 1 e poi fusi con la regola della somma pesata (Nanni et alii, 2012). Per ciascuna rappresentazione matriciale  $rm$  (ID/SC/HF) i risultati del confronto diretto e dei relativi descrittori di tessitura LPQ è combinato usando la seguente formula:  $6 \times rm + LPQ3_{rm} + LPQ5_{rm}$ .

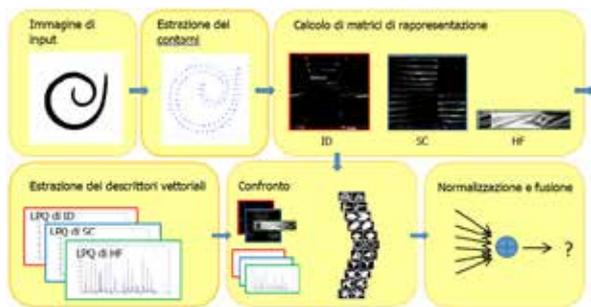


Fig. 3 - Rappresentazione schematica del metodo di confronto.

## RISULTATI SPERIMENTALI

Il metodo proposto in questo lavoro è stato testato su due problemi di classificazione relativi ai due dataset descritto in precedenza: classificazione dei motivi dei vasi Kamares (KAM); riconoscimento dei cocci dei vasi trovati nel porto di Classe.

Nel caso del dataset Kamares le immagini ci sono state fornite dagli autori di (Stanco et alii 2011) e abbiamo utilizzato solo i motivi per i quali erano presenti almeno 2 esempi (in totale 61 immagini sono state utilizzate).

Nel caso del secondo dataset prendiamo in considerazione due diverse problematiche di classificazione:

a) VAS\_1, riguarda la capacità di distinguere a quale zona dell'anfora appartiene il reperto: 'fondo', 'orlo', 'orlo+collo' e 'orlo+collo+ansa';

b) VAS\_2, riguarda la capacità di determinare a quale tipologia di anfora appartiene il reperto. Considerato che in alcuni casi vi sono pochi esempi per tipologia, si è stati costretti ad eliminare i casi rari, e si considerano solo gli elementi 'fondo' e 'orlo'. In più, vengono considerate le tipologie con almeno 5 esempi così da avere in tutto 8 tipologie. Per tutti e tre i problemi di classificazione è stato utilizzato lo stesso protocollo di test: per ogni immagine viene calcolata la distanza con tutte le altre immagini, poi l'immagine viene classificata in base all'immagine del training più vicina.

Il protocollo utilizzato è simile a quello usato in (Stanco et alii 2011; Guarnera et alii 2010) dove è stato ottenuto un risultato del 75% di accuratezza di classificazione. Il sistema proposto in questo lavoro (denominato ENS in tabella 1) ottiene le migliori performance indipendentemente dal dataset e dal protocollo di testing. Gli altri risultati riportati in tabella 1 riguardano sistemi di classificazione più semplici, basati sull'uso di un solo tipo di descrittore (ID, SC, HF) E' molto interessante notare che il metodo ID ottiene performance peggiori di SC in quasi tutti i problemi, cosa che non accade in altre problematiche (Nanni et alii 2012).

Nel caso VASI\_1 e KAM i risultati ottenuti dal metodo completo ENS sono molto buoni e incoraggianti, anche rispetto all'utilizzo del solo metodo ID: ottenere un'accuratezza di classificazione prossima al ~90% usando dataset di dimensioni non elevate, lascia prevedere l'acquisizione di nuovi dati non potrà che fare aumentare le performance.

		Dataset		
		VASI_1	VASI_2	KAM
Metodo	IDSC	87.8%	64.8%	86.9%
	SC	89.5%	68.2%	86.9%
	HF	88.9%	68.2%	88.5%
	ENS	91.2%	68.9%	88.5%

Tabella 1 - Risultati ottenuti dai metodi testati in questo lavoro.

## CONCLUSIONI

Il sistema che abbiamo proposto può soddisfare l'obiettivo che volevamo raggiungere, cioè un sistema automatico in grado di velocizzare il riconoscimento di forme in diverse problematiche.

Il sistema per ora è stato testato utilizzando informazioni estratte dalla forma dell'oggetto, in realtà si dovrebbe studiare anche la presenza di ulteriori forme all'interno dell'oggetto stesso, un esempio è fornito dai reperti che presentano 'orlo', 'collo' e 'maniglia', in cui al loro interno, il sistema, rileva un foro.

Nonostante questo particolare, si può affermare che i risultati sono abbastanza buoni, tanto da far pensare che la fase di riconoscimento del motivo, della classe ceramica o della tipologia possa essere automatizzata. Velocizzando così il lavoro necessario ad affrontare la catalogazione e l'organizzazione dei reperti e facilitare la creazione di un database per i reperti trovati in ogni scavo.

## RINGRAZIAMENTI

Desideriamo ringraziare per la preziosa collaborazione Filippo Stanco per averci fornito il dataset delle immagini dei motivi delle ceramiche Kamares.

## BIBLIOGRAFIA

- Belongie, S., Malik, J., Puzicha, J., 2002. Shape matching and object recognition using shape contexts. *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (24), 509-522.
- Bonifay M., *Etudes sur la céramique romaine tardive d'Afrique*, Archaeopress, 2004
- Boschi F., *L'idrografia*, in *Classe. Indagini sul potenziale archeologico di una città scomparsa*, a cura di AUGENTI A., 2011, pp. 199-204.
- Guarnera, G. C., Stanco, F., Tanasi, D., & Gallo, G. (2010). Classification of decorative patterns in Kamares pottery. In M. Samuelcic (a cura di), *Proceedings of SCCG 26th Spring Conference on Computer Graphics* (pp. 20-23).
- Keay S. J., *Late roman amphorae in the western Mediterranean: a typology and economy study, the Catalan evidence*, 1984.
- Krassimira I., *Content-Based Image Retrieval in Digital Libraries of Art Images Utilizing Colour Semantics*. Proceeding of: *Research and Advanced Technology for Digital Libraries - International Conference on Theory and Practice of Digital Libraries, TPDL 2011, Berlin, Germany, September 26-28, 2011*.
- Ling, H., Jacobs, D.W., 2007. Shape classification using the inner-distance. *IEEE Trans. Pattern Anal. Mach. Intell.* 29 (2), 286-299.
- Lumini A. and D. Maio, *A shape similarity approach for searching images in a large database of works of art*, in *proceedings 2nd International Congress on "Science and Technology for the Safeguard of Cultural Heritage in the Mediterranean Basin"*, Paris (France), July 1999.
- Lumini A., *Fruizione di beni artistici in formato digitale: un approccio al repertorio di immagini per similarità*, in *proceedings Workshop "Intelligenza artificiale per i Beni Culturali" (99)*, Bologna (Italy), pp.99-108, September 1999
- Nanni L., Brahmam S. and Lumini A., *Local phase quantization descriptor for improving shape retrieval/classification*, *Pattern Recognition Letters*, vol.33, no.16, 2012.
- Ojanstivu, V., Heikkila, J., 2008. Blur insensitive texture classification using local phase quantization. In: *ICISP2008*.
- Stanco F., Tanasi D., Guarnera C., Gallo G., (2011) *Automatic classification of decorative patterns in the Minoan pottery of Kamares style*, in C. Papaodysseus (a cura di), *Pattern Recognition and Signal Processing in Archeometry: Mathematical and Computational Solutions for Archeology*, IGI Global.
- Stanco F., Tanasi D., (2011) *Experiencing the Past. Computer Graphics in Archaeology*, in F. Stanco, S. Battiato, G. Gallo, (a cura di), *Digital Imaging for Cultural Heritage*, CRC Press.
- Wang, J., Bai, X., You, X., Liu, W., Latecki, L.J., 2012. Shape matching and classification using height functions. *Pattern Recogn. Lett.* 33 (2), 134-143

## ABSTRACT

THE CLASSIFICATION OF SHAPE (E.G. PIECE OF POTTERIES, MOTIF IDENTIFICATION, HIEROGLYPHS CLASSIFICATION...) IS A TIME CONSUMING TASK FOR HUMAN BEING, THEREFORE THE OBJECTIVE TO DEVELOP AN AUTOMATED SYSTEM FOR SHAPE CLASSIFICATION IS WELCOMED FOR MANY ARCHAEOLOGISTS, FREING THEM FROM TEDIOUS EFFORTS. MOREOVER, DUE TO THE LACK OF HUMAN EXPERTS SEVERAL ARTIFACTS ARE NOT STUDIED, THIS PROBLEM COULD BE ALLEVATED USING AUTOMATED SYSTEMS. A MAIN PROBLEM OF SEVERAL PROPOSED SHAPE CLASSIFIERS IS THAT THEY NEED A PARAMETERS TUNING IN EACH DIFFERENT PROBLEM. WE SHOW THAT OUR SYSTEM BASED ON AN ENSEMBLE OF APPROACHES IS ABLE TO WORK IN DIFFERENT APPLICATIONS EXACTLY WITH THE SAME PARAMETERS. THIS IS A VERY USEFUL FOR PRACTITIONERS, WITH LITTLE KNOWLEDGE OF PATTERN RECOGNITION, THAT NEED A QUICK METHOD FOR THEIR SHAPE CLASSIFICATION PROBLEM. TWO CASE STUDIES ARE HERE REPORTED: AUTOMATIC CLASSIFICATION OF THE SEVERAL DIVERSE MOTIFS THAT ARE USED IN THE POTTERY OF KAMARES STYLE (CRETAN CULTURAL PRODUCTION BETWEEN THE FIRST HALF OF THE 2ND MILLENNIUM BC); PIECES OF POTTERY CLASSIFICATION FOUND IN THE AREA OF THE ANCIENT PORT CITY OF CLASS (ITALY). THE MATLAB CODE OF THE PROPOSED SYSTEM WILL BE AVAILABLE AT [HTTP://WWW.DEI.UNIPD.IT/WDYN/?IDSEZIONE=3314&IDGRUPPO\\_PASS=124&PREVIEW=AS](http://www.dei.unipd.it/wdyn/?IDSEZIONE=3314&IDGRUPPO_PASS=124&PREVIEW=AS) AS BASELINE FOR OTHER RESEARCHERS.

## PAROLE CHIAVE

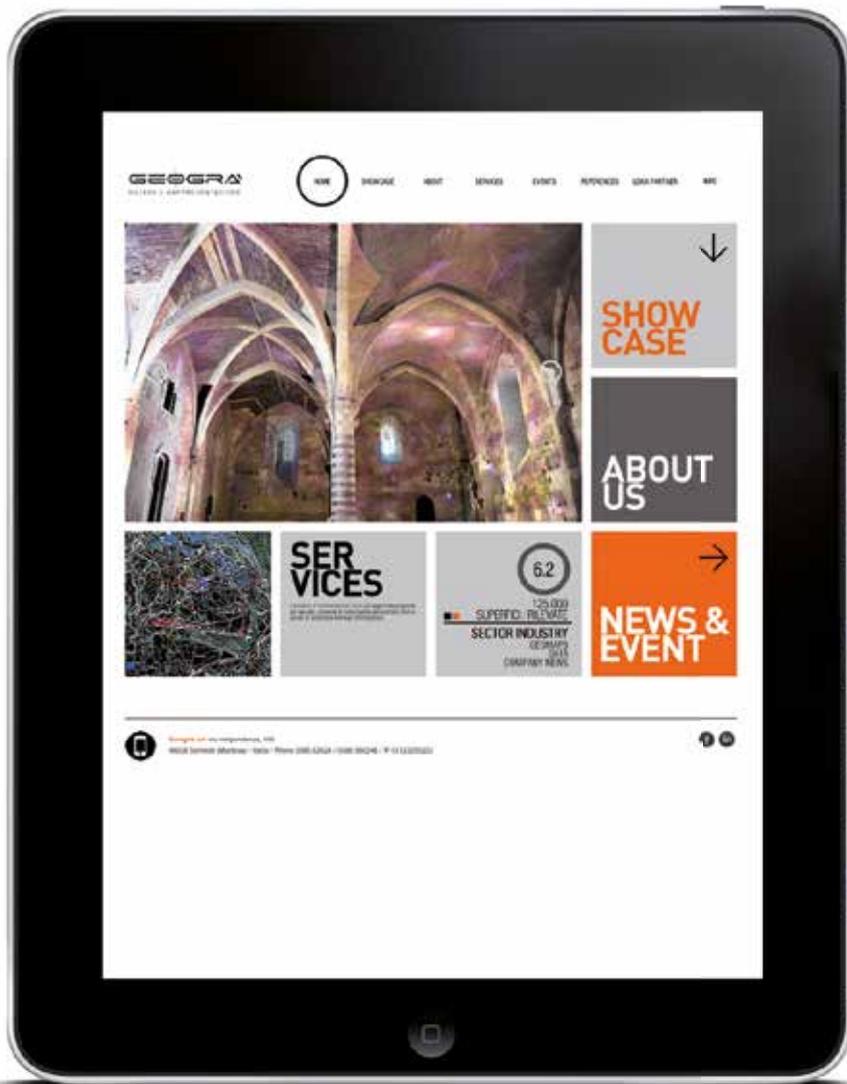
COMPUTER VISION; PATTERN RECOGNITION; DATABASE; DESCRITTORI DI FORMA; VASI; MOTIVI DECORATIVI

## AUTORI

LORIS NANNI, [loris.nanni@unipd.it](mailto:loris.nanni@unipd.it)  
DEPARTMENT OF INFORMATION ENGINEERING, VIA GRADENIGO 6/A, 35131- PADOVA ITALY

ALESSANDRA LUMINI  
DISI, UNIVERSITY OF BOLOGNA, VIA VENEZIA 52, 47023 CESENA ITALY

MONICA TOSCHI  
DIPARTIMENTO DI BENI CULTURALI — VIA DEGLI ARIANI, 1 48121 - RAVENNA - ITALIA



# OBSERVE WITH US

Siamo specializzati nel rilievo metrico in ambito architettonico, archeologico, infrastrutturale, impiantistico, batimetrico e topografico. Il nostro sistema di prendere possesso degli spazi che ci circondano, è essenzialmente un modo visuale e percettivo in cui spazio e tempo vengono restituiti come sistemi continui e iconici.



**GEOGRA**

via indipendenza, 106 / 46028 / sermide (mn) / italy  
phone +39.0386.62628 / info@geogra.it

[www.geogra.it](http://www.geogra.it)