NEW THESAURUS QUALITIES OF ARBOR

Abstract

ARBOR is a formal language developed for describing hierarchically organized (i.e. tree-structured) data-objects like f.i. the textual representation of the mostly complex research objects of classical archaeology or history of art (e.g. works of architecture like ancient temples or churches or of figural arts like sculptures or paintings or any combination of these or other object types). ASU (Arbor—Set—Up) is a program which translates from an "ARBOR-text" containing a thesaurus and a set of object descriptions (i.e. "documents") into a computer-readable database. Inside the latter the program ART (Arbor—ReTrieval) allows the interactive and tree-structure (or-set-and-subset, etc.) specific research (Eisner 1989a, 1989b). Until now the working programs are DOS-based versions for PC's and compatibles. In 1988 ARBOR was integrated in the GMD-IPSI-project PAVE (Publication And Visualisation Environment) and in the future ARBOR will play the role of the textual module in databases storing visual information f.i. about archaeological and art-history objects. Because this objective breaks the limits of DOS ASU and ART actually are ported to UNIX.

Treating descriptions of real objects very often there is the desire to address them also under broader terms, f.i. temples and churches under "religious buildings" or even more generally as "architecture" or sculptures and paintings under "visual-art-objects". In earlier program-versions ARBOR-thesaurus-functions were limited to a control of the terminology used in the single object descriptions (or "documents") of an "ARBOR-text". For later retrieval using also broader terms there was no other possibility than to integrate the latter in the respective descriptions and the result was a certain terminological redundancy. The newest version of ASU additionally to normal "ARBOR-descriptors" allows the use of three new kinds of thesaurus elements, namely "pointers", "superpointers" and "superdescriptors". Working with ART they cause the reference to single descriptors or to sets of them. Inside the thesaurus the four types of elements can be combined freely and even nested giving the possibility to represent also complex generic relations or terminological hierarchies. These additional qualities of ARBOR lead to simpler (i.e. less redundant) object descriptions and to more retrieval power in order to ease setting up and using pictorial databases.

1. Hierarchical relations in thesauri

In thesaurus science two hierarchical relations between terms are to be distinguished: the "partitive" and the "generic" one. In the partitive relation the subordinate terms denominate parts of the superordinate term (Neveling, Wersig 1975, 100). That may be shown by an example. One of the best known types of the ancient Greek temple is the so called "peripteros" which consists of a "peristasis" or exterior colonnade, the "pteron" or aisle behind the colonnade and the "naos" or core of the peripteros. The naos for its part is composed of the "pronaos" or porch, the "cella" or main room and the "opisthodomos" or rear room (Fig. 1).

The generic relation between terms is quite different: the subordinate terms have the same characteristics as the superordinate one and at least one more (Neveling, Wersig 1975, 99). Let us give one example. The Roman architect Vitruvius in his famous work *de architectura libri decem* 68, 25 speaking about temples distinguishes other six types besides the above mentioned peripteros, and that are the "templum in antis", the "prostylos", the "amphiprostylos" the "pseudodipteros", the "dipteros" and the "hypaethros"(Fig. 2)[1]. Thus these seven terms denominate reifications of the more general term "temple" which means only cultic building in any form.

2. Data hierarchies in ARBOR

The partitive relation is useful not only for giving structure to thesauri but also for the conceptual description of the mostly complex objects of the classical archaeology and history of art. That is the case not only with temples or other works of architecture, like ancient theaters, aqueducts or sepulchral monuments, but also with statues, reliefs or vases, or more in general with objects bearing f.i. pictorial representations. It is just the textual documentation of this kind of objects that the formal language ARBOR was developed for. It allows both the description and the evaluation of objects differentiated by their respective detail levels.

The use of a thesaurus was possible already in earlier versions of ARBOR, but it was active only in the moment of setting up a database. In the last version however the thesaurus has also retrieval functions.

An ARBOR-database or "ARBOR-text" (or "ARBOR-file") normally consists of a series of "documents" which will describe different objects of classical

[1] The classification given by Vitruvius shows some minor differences to the typology as it is usual today, mainly because the hypaethros does not represent a real ground-plan typus but only a temple with an uncovered cella.

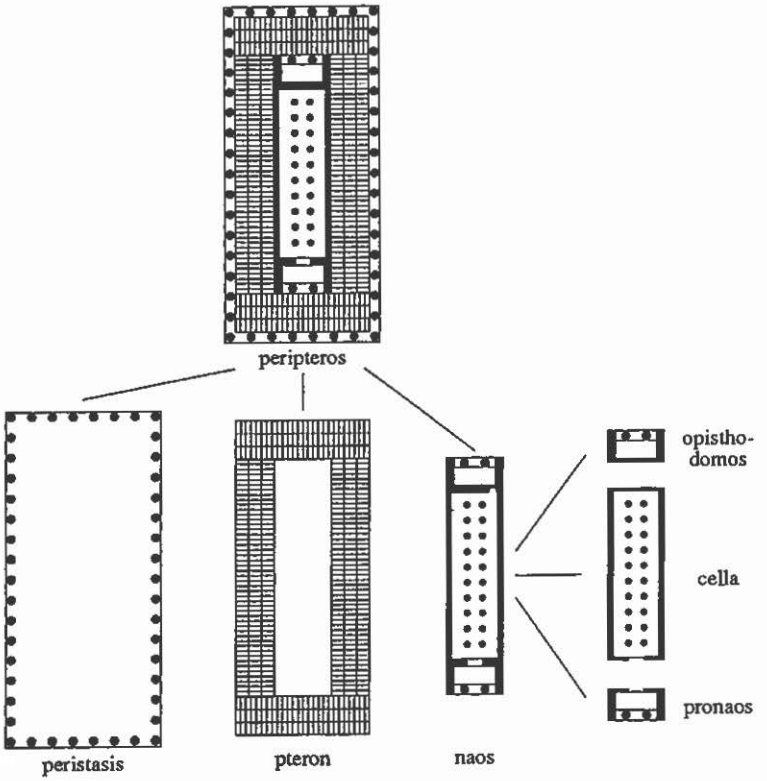


Fig. 1 — A "peripteros" and its single parts.

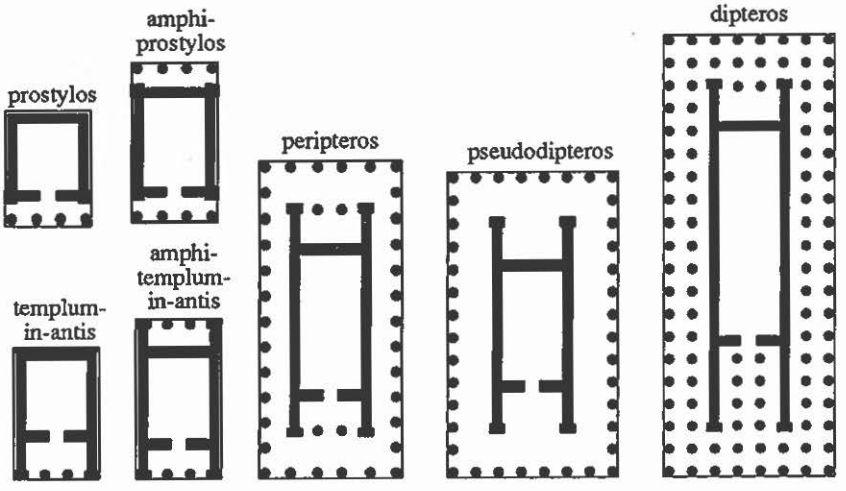


Fig. 2 — The different classes of ancient Greek temples.

archaeology or history of art. Inside every document there are to be distinguished "descriptors" and "commentaries", but later only the first ones will be retrievable. The first document of an ARBOR-text always will be the thesaurus. Normally it does not remain empty and contains descriptors which, as we indicated, serve to control the terminology used in the subsequent documents. That is to say that every descriptor used in a document must have appeared at least once before in the thesaurus. The begin and the end of the thesaurus are marked by the respective labels ("$THS + " and "$THS-"):

```
$THS +
  [thesaurus] ...
$THS-
*Priene, *temple of *Athena-Polias (by architect *Pytheos), *erected = -340..-325 (means
340-325 B.C.), *peripteros, *style:ionic
- *crepidoma
- *peristasis
- *pteron
- *naos
-- *pronaos
-- *cella
-- *opisthodomos
   ...
[next document]
...
```

Contrary to the thesaurus the following documents do not contain lists of allowed terms but descriptions of real objects of classical archaeology or history of art. Every document must have a "head" whose descriptors will describe the object as a whole, like f.i. name, location or chronological data. If the object consists of clearly discernible subsets (or subsets of subsets, etc.) like f.i. peristasis, pteron or naos of a peripteros (or pronaos, cella and opisthodomos of its naos) there can be added correspondent "subdocuments" which on their part will be composed of descriptors and commentaries. Every subdocument has to begin with a "contextor" which corresponding to the hierarchical dependence directly of the header or of an antecedent subdocument consists of one or more hyphens "-"). But these contextors are allowed only in object-describing documents and not in the thesaurus, because the latter has not at all the function of representing partitive relations between the subsets of complex objects but at the contrary term-lists which normally have a linear structure.

3. Retrieval with thesaurus functions

Like other information systems ARBOR was designed starting from an imaginary user query. In the hitherto existing versions the thesaurus was limited to the terminological control during the setting up of a database. In the last version

this control is extended also to the user query. What in the first moment seems to be nothing more than a mere technical detail de facto gives a new possibility which is not to be undervalued, i.e. to substitute descriptors during the query.

In this regard three cases were taken into consideration:

1. Replacing one query term by one descriptor like it will be necessary with synonyma. Let us give an example which has to do with a database containing data about the photographic archive of the author. In the thesaurus appears the descriptor "Sardes" meaning the ancient Anatolian town of Lydian origin. In order to give the user the possibility to ask also for "Sardis" which is the other very common version of the town-name a "pointer" can be included in the thesaurus which points to the regarding descriptor. For this purpose the item "Sardis" must be marked not as usual for a descriptor with an asterisk ("*") but with an up-arrow ("^") and be placed before the pointed target, i.e. "Sardes", in order to put it in the following way:

   ```
   ^ Sardis *Sardes
   ```

2. Replacing one query term by more than one descriptor, like in the following example taken from the same database. The user may search for a term which for avoiding redundancy is not allowed in object descriptions, f.i. "Ionia" (together with "Lydia" another of the main countries of Western Anatolia in Antiquity). In this case you can admit "Ionia" as a "super-pointer" in the thesaurus which now points not to a single descriptor but to a set of them enclosed by a pair of curly braces like when writing:

   ```
   $THS +
   ...
   ^Ionia {
     *Belevi
     *Clarus
     *Didyma
     *Ephesus
     *Miletus
     *Phokaia
     *Priene
     *Teos
        }
   ...
   $THS-
   ```

   That is a concrete example for the generic relation as described at the beginning of this paper. Another one also belonging to the archive data will be the relation

   ```
   ^religious-architecture    {
     *synagogue
     *temple
     *church
                              }
   ```

3. Enhancing one descriptor by a series of other descriptors like in

```
*temple    {
  *prostylos
  *amphiprostylos
  *templum-in-antis
  *amphi-templum-in-antis
  *peripteros
  *pseudodipteros
  *dipteros
  ...
            }
```

in order to come to a "super-descriptor". At the first glance this seems to be a weakened form of the properly speaking unambiguous generic relation. But it makes sense in a science like archaeology where the desired objectivity of the scientific statement very often is put in question by the defective state of conservation or disclosure of the respective treated material objects. The foundations of a methodically excavated and measured ancient temple are easily to be equated with one of the cited temple types. A less well analyzed example however may be classified only as "temple" in general but not as one of its concrete incarnations. So the user searching for "temple" if required will be able to find all items having anything to do with this descriptor independently of the degree of distinctiveness of the single finds in the respective ARBOR-text.

At the end it may be pointed to the possibility to combine pointers, super-pointers and super-descriptors with each other which enables to represent more complicated synonym and generic relations.

MICHAEL EISNER
Integrated Publication and Information System Institute
Gesellschaft für Mathematik und Datenverarbeitung Darmstadt

BIBLIOGRAPHY

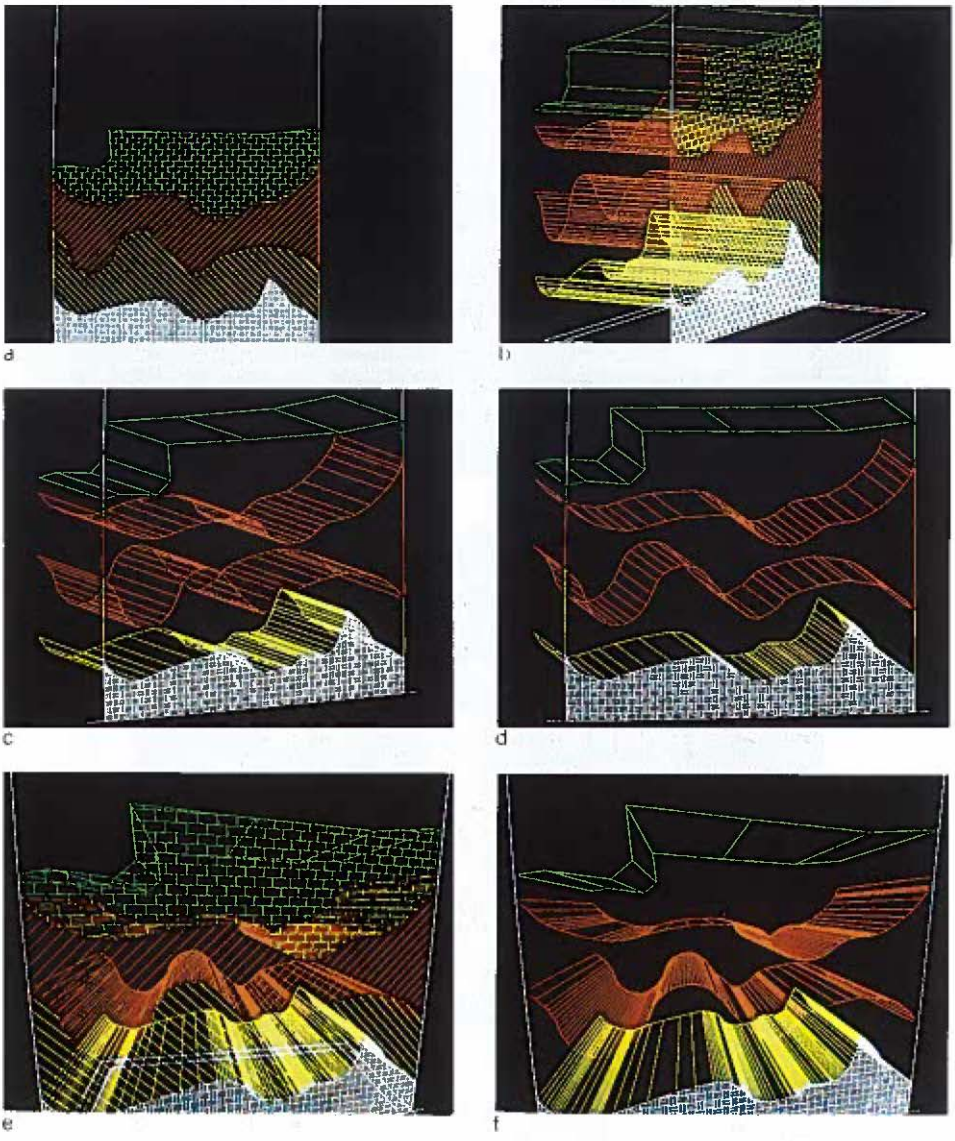
EISNER M. 1989a, *ARBOR. Eine Sprache zur Beschreibung und ein Programmpaket zur Verarbeitung hierarchischer Datenobjekte der klassischen Archäologie und Kunstgeschichte*, GMD-Studien Nr. 159, Birlinghoven, 64 pp.

EISNER M. 1989b, *The ARBOR Information System for Classical Archaeology and History of Art*, « International Classification », 16, pp.141-145.

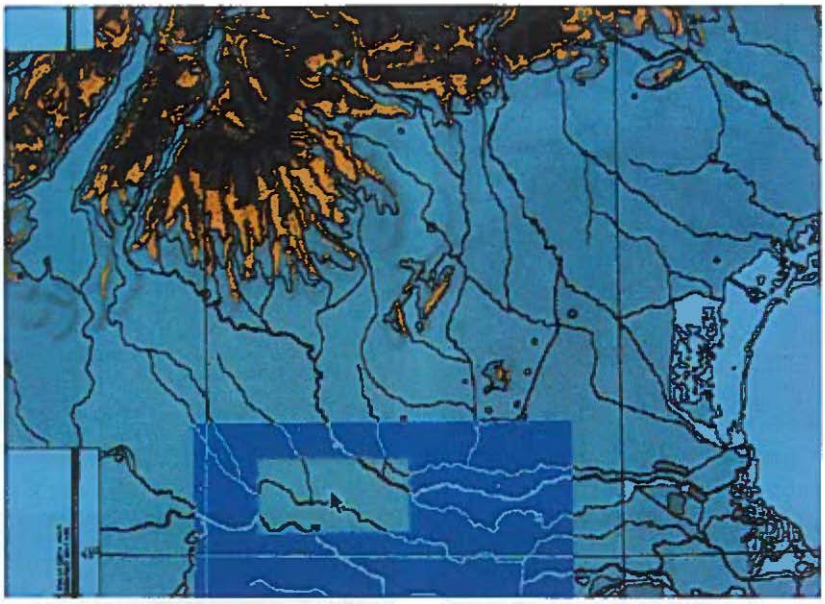
NEVELING U., WERSIG G. 1975, *Terminologie in der Information und Dokumentation*, München.

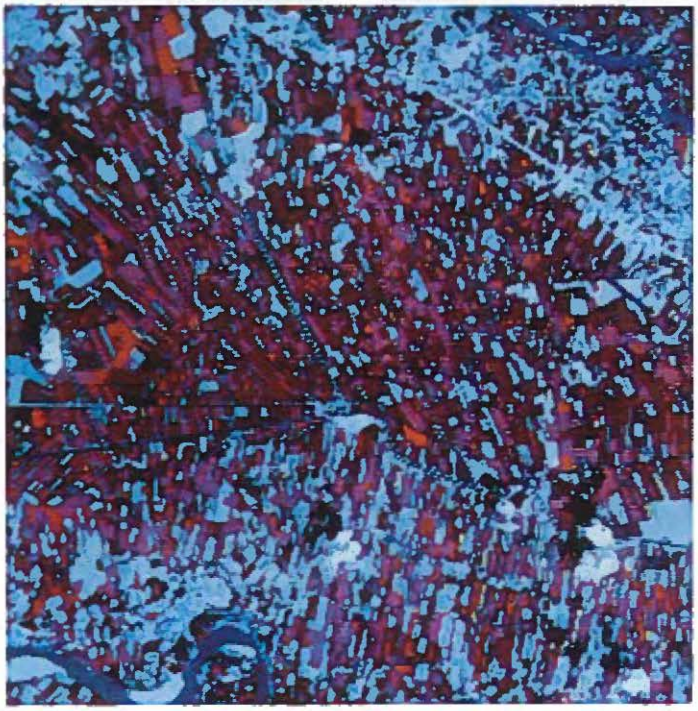
ABSTRACT

The Author describes new additional qualities of the program ARBOR which allow a less redundant object description and facilitate the setting up and the use of pictorial databases.

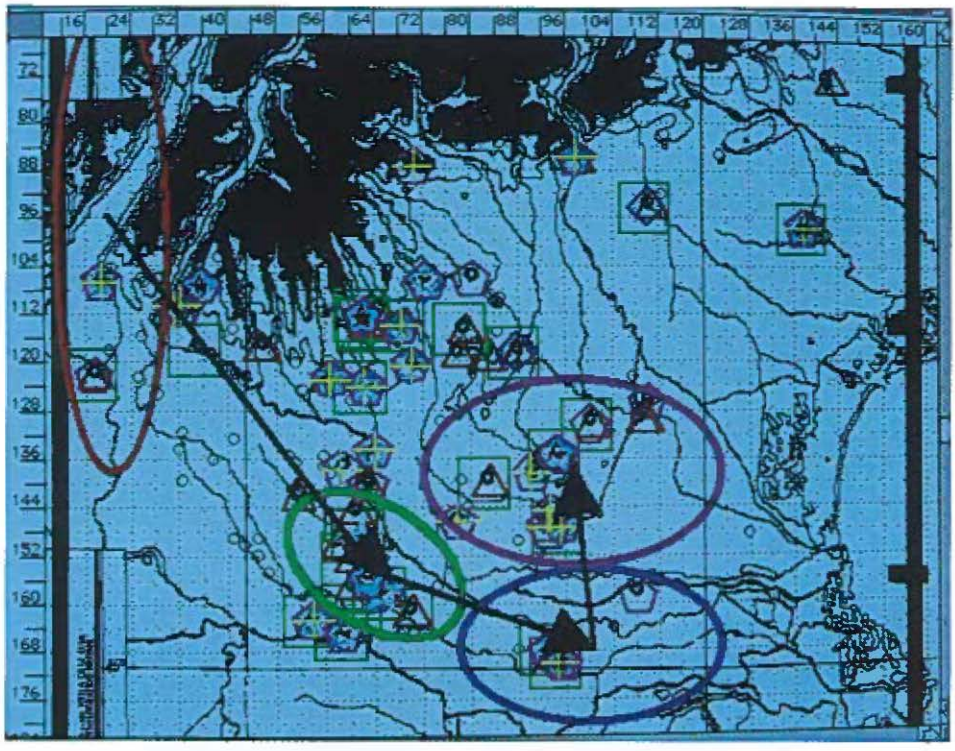


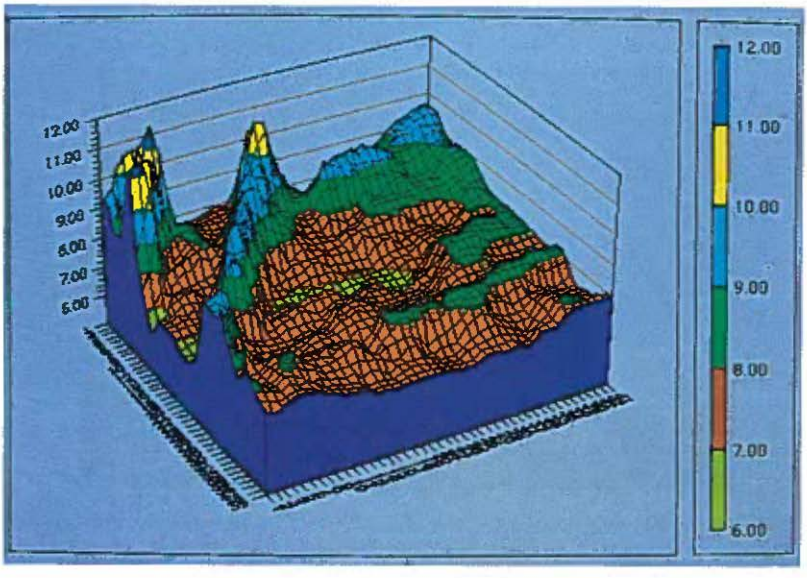
Tav. I — Esemplificazione di realizzazione grafica computerizzata di una sezione archeologica con Autocad 10, in visione tridimensionale. a, fronte della sezione; b, c, d, punti di vista diversi della stessa sezione con gli strati visti in profondità e con evidenziazione delle interfacce; e, vista frontale e in profondità; f, vista frontale delle interfacce.



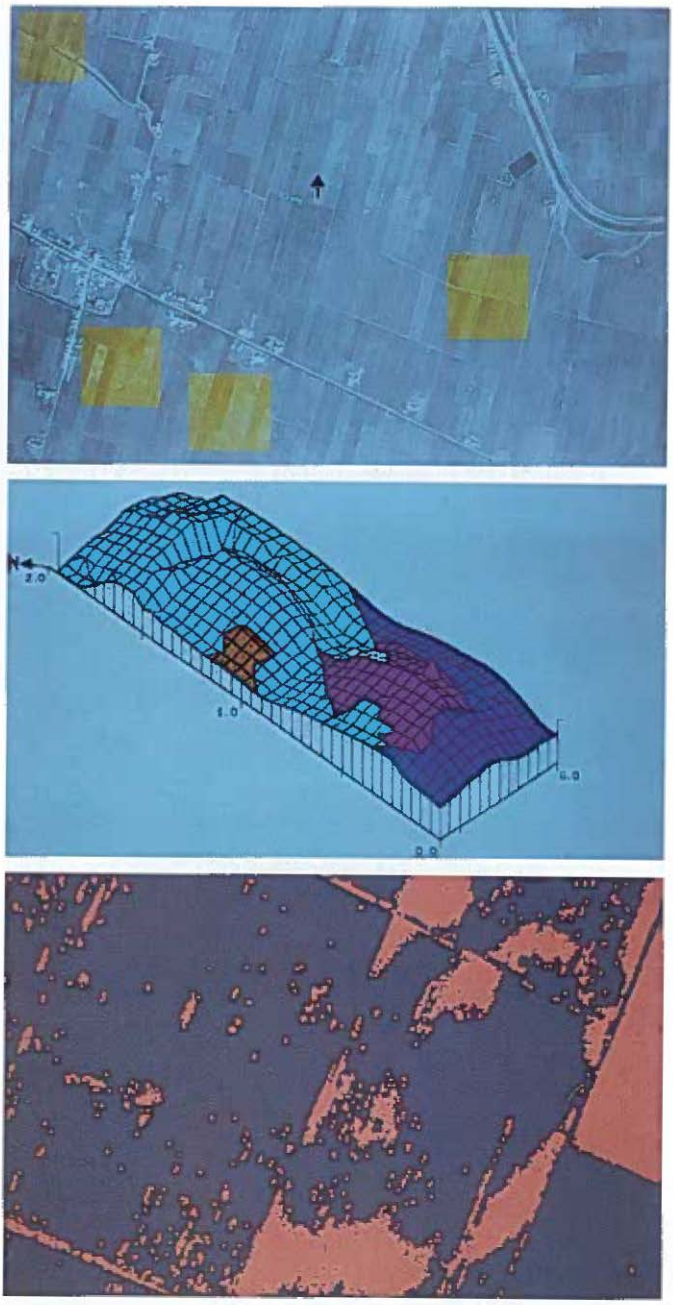


Tav. II — a) Mappa del Veneto con l'indicazione dell'universo campionario e dell'areale operativo del "Progetto Alto-Medio Polesine-Basso Veronese".
b) Foto da satellite (satellite: Spot 1; sensore: HRV1; data: 25 luglio 1986; scala originale 1:47.0000 circa; risoluzione: 20 x 20 m.; composizione: falso colore di Spot multispettrale; bande: infrarosso, rosso, verde su rosso, verde, blu; orbita: k62J2597; angolo di ripresa: 4.9 Ovest) sull'attuale areale operativo del progetto fra Po (cfr. in basso l'ansa di Bergantino-Castelnovo Bariano (Ro), e l'Adige (cfr. in alto a destra l'ansa di Carpi di Villabartolomea — VR) solcato nel mezzo dal corso del Tartaro (elaborazione e dati tecnici di C. Giaggio — Regione Veneto-Dipartimento per le Foreste e l'Economia Montana).

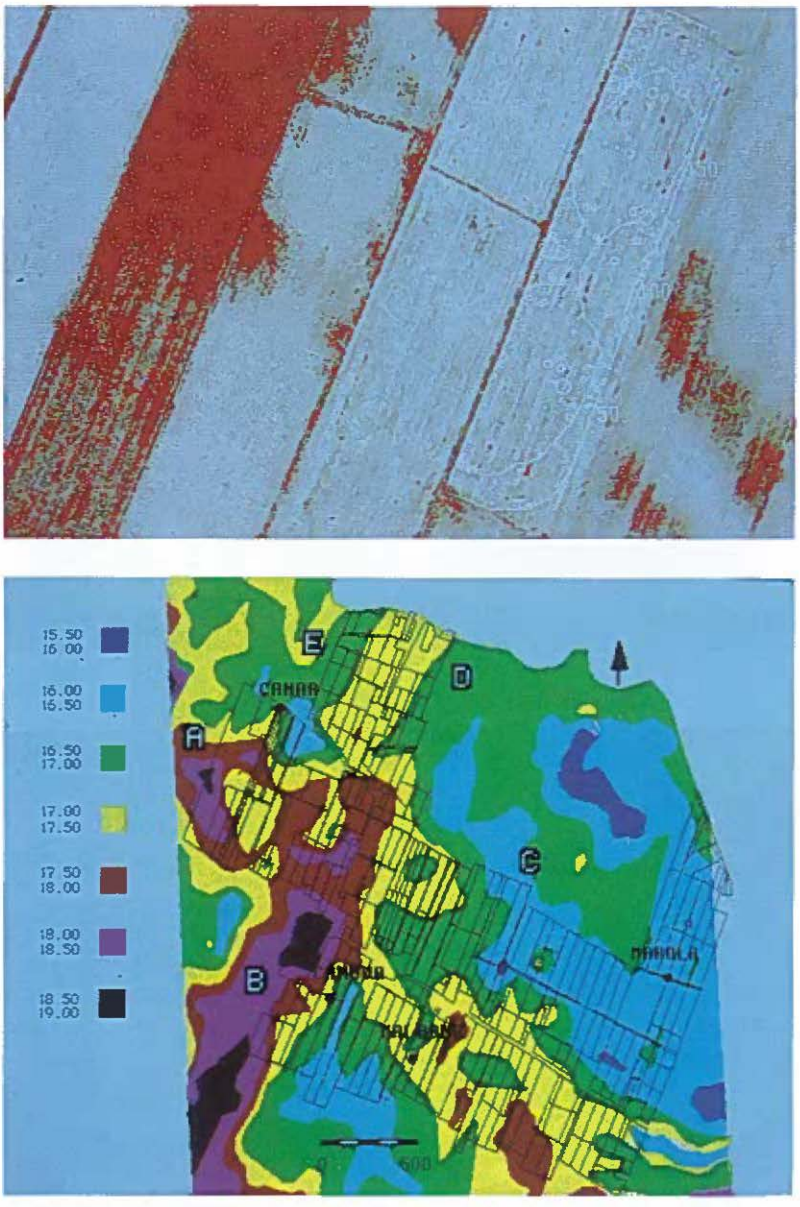




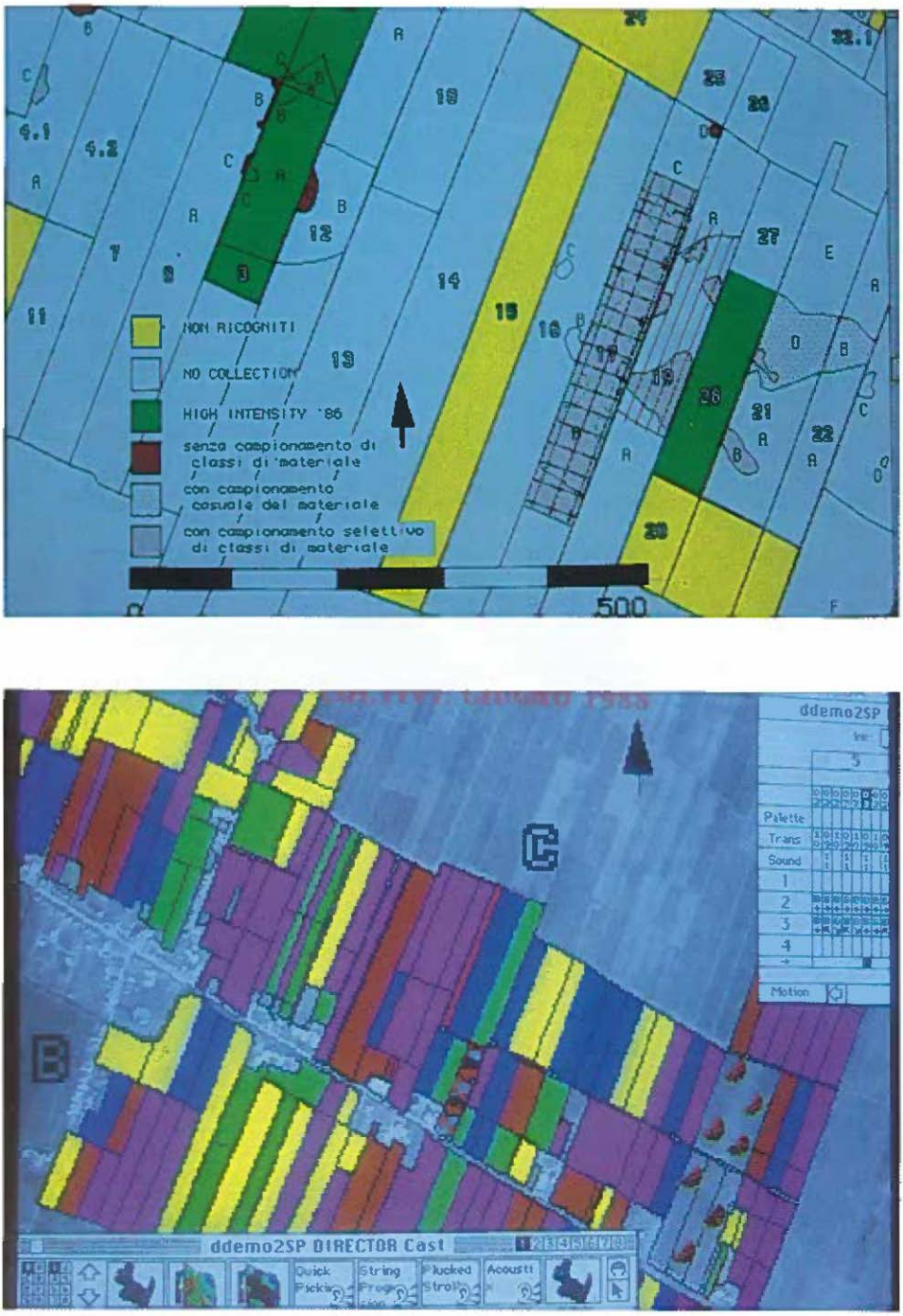
Tav. III — a) Mappa archeologica plurifase del Veneto fra Bronzo Recente Evoluto e Inizio dell'età del Ferro: Fasi 1-6, contraddistinte dal simbolo e dal colore. Fase 1 = triangolo rosso; Fase 2 = rettangolo verde; Fase 3 = rombo blu; Fase 4 = stella celeste; Fase 5 = pentagono rosa; Fase 6 = croce gialla. Il flusso diacronico delle "core areas" (insiemi col colore pertinente alla fase-climax) è evidenziato dall'andamento delle frecce (cfr. BICEGO 1990).
b) Rappresentazione tridimensionale del microrilievo dell'areale del sito di Fabbrica dei Soci (cfr. Tavv. VIIb-VIIIa).



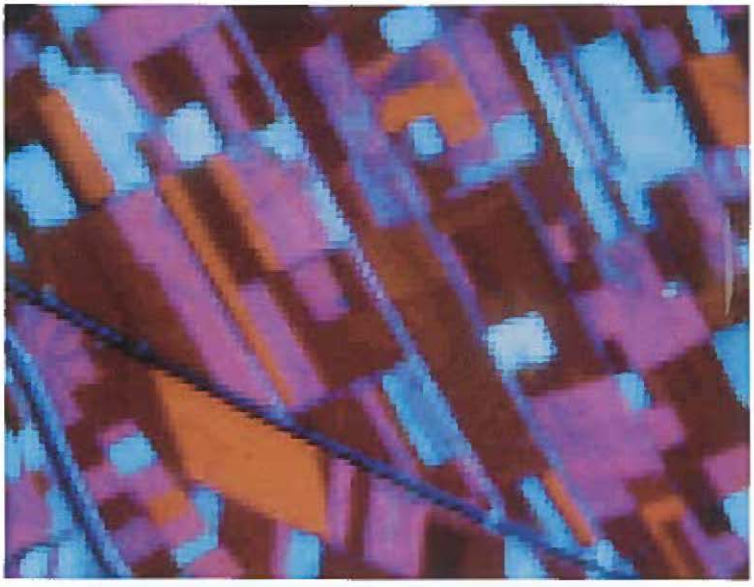
Tav. IV — a) S. Pietro Polesine (Foto aerea C.G.R. — Parma — 12/5032 — scala originale del fotogramma di 1:16600 circa; Conc. S.M.A. n. 1206 del 28/11/89). Le quattro aree selezionate corrispondono (in sequenza da Ovest ad Est) ai siti principali di Canar, Canova, Malbona, Marola (cfr. testo).
b) Canova — campo 3: rappresentazione tri-dimensionale dei microrilievo in relazione ai contesti di superficie.
c) Immagine digitalizzata del sito di Canova, dopo una sequenza di processi di "trattamento di immagine" con il programma Optilab (cfr. testo): equalizzazione, "thresholding", morfologia primaria (erosione e dilatazione) e avanzata (P—Opening). (Da foto aerea C.G.R. — Parma — 12/5032 — zoom da scala originale del fotogramma di 1:16600 circa; Conc. S.M.A. n. 1206 del 28/11/89).

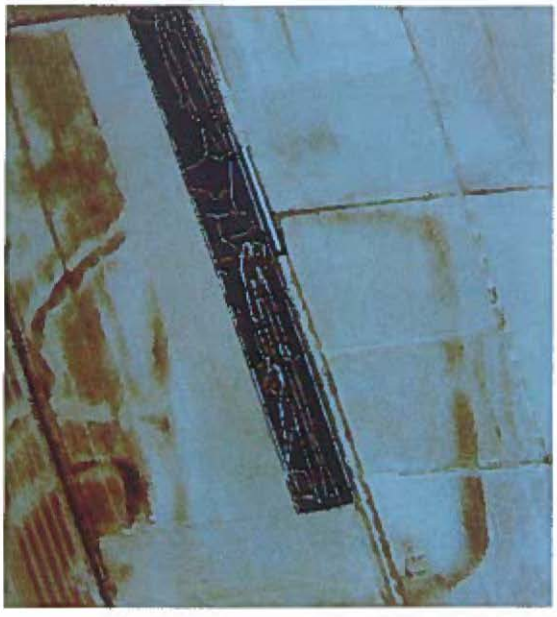


Tav. V — a) Immagine digitalizzata del sito di Malbona: sequenza di processi di "trattamento di immagine" con il programma Optilab (cfr. testo): equalizzazione, rimappatura di 16 livelli di grigio su scala cromatica "temperatura dell'acqua", "thresholding". Aree più scure = terreni scoperti a tessitura sabbioso-limosa; aree più chiare = terreni scoperti a tessitura limoso-argillosa e campi coltivati; bianco = mappa composita del microrilievo (curve di livello ai 50 cm.) e della frequenza assoluta normalizzata del numero di frammenti ceramici (simbolo = cerchio di ampiezza proporzionale alla frequenza) sovrapposta in trasparenza. I layers al tratto, ottenuti dal pacchetto statistico Systat (SYSTAT, inc.), sono stati prima sovrapposti all'immagine aerea, usando le sofisticate funzioni di pasting di Digital Darkroom: cfr. testo). (Da foto aerea C.G.R. — Parma — 12/5032 — zoom da scala originale del fotogramma di 1:16600 circa; Conc. S.M.A. n. 1206 del 28/11/89; cfr. riferimento di scala metrica sul lato ovest del campo).
b) Area del survey di "media intensità" (S. Pietro Polesine 1987; 1988): mappa composita del microrilievo (isoipse ai 50 cm.) e del perimetro dei campi (elaborazione grafica computerizzata di F. Cafiero).

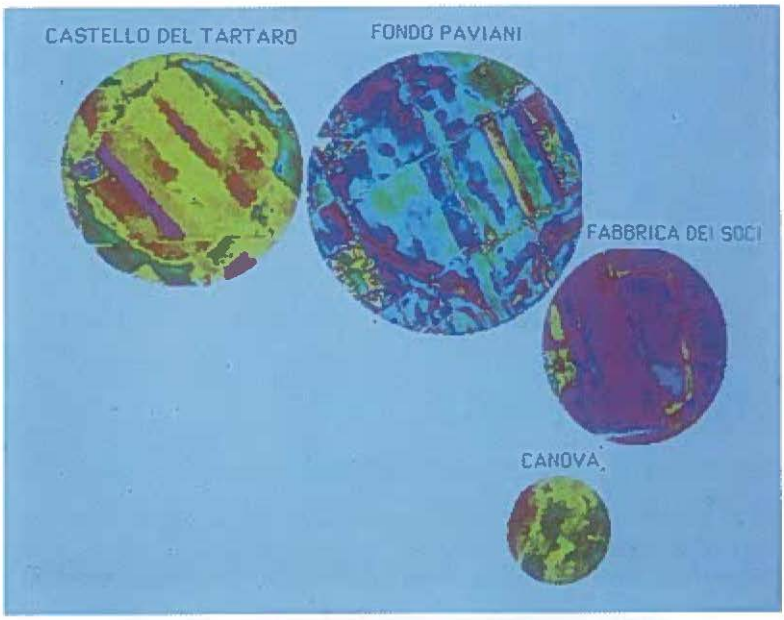


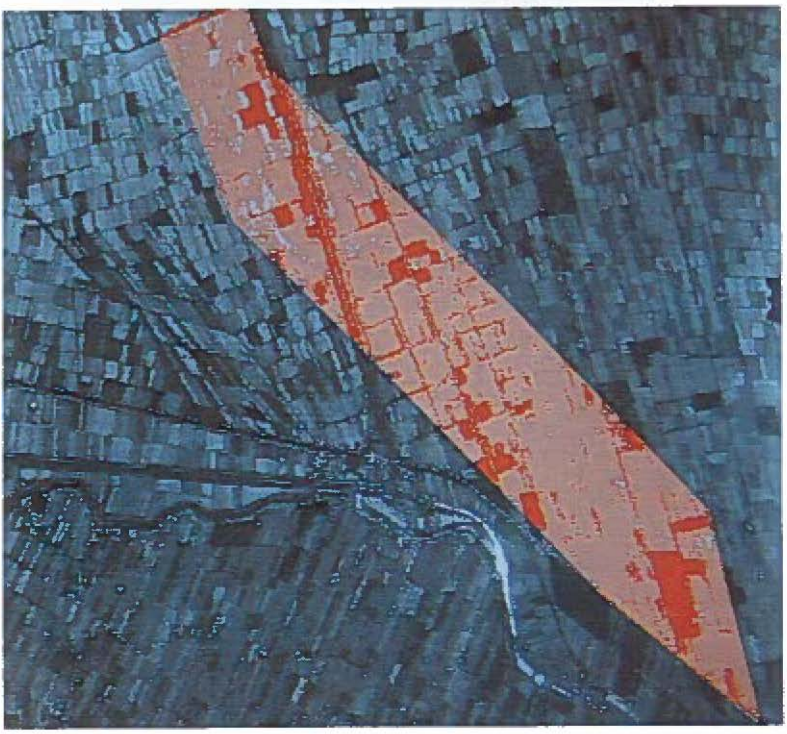
Tav. VI — a) Survey di "media intensità" (S. Pietro Polesine 1987-1988) — locc. Malbona e Canova: mappa delle strategie di raccolta (elaborazione computerizzata di A. Malgarise).
b) Dettaglio sull'areale del survey di "media intensità" (S. Pietro Polesine 1987-1988): mappa di coltivi (Giugno 1988) sovrapposta in "opaco" al layer dell'immagine aerea (cfr. Tav. IVb) (da una sequenza multimediale ottenuta con il programma Director: cfr. testo).





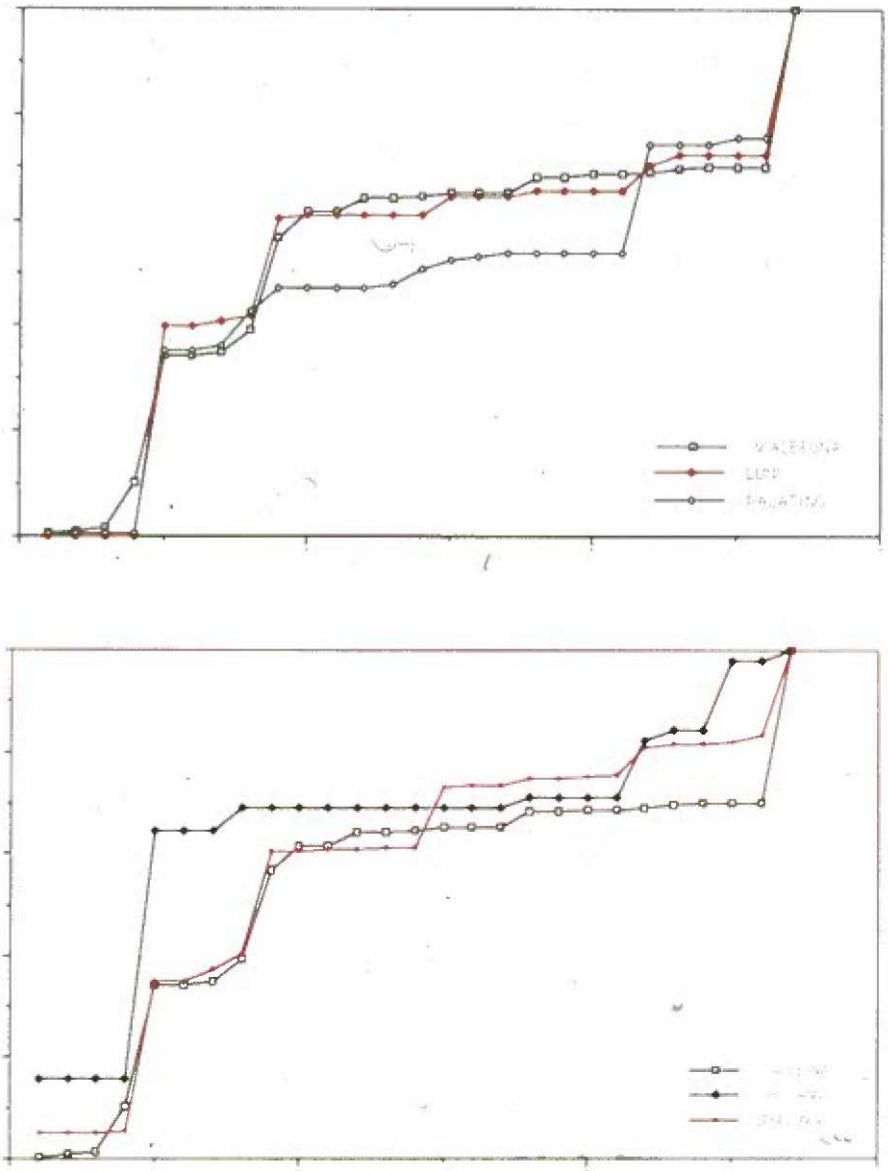
Tav. VII — a) Foto da satellite (satellite: Spot 1; sensore: HRV1; data: 25 luglio 1986; scala originale 1:47.0000 circa; risoluzione: 20 x 20 m.; composizione: falso colore di Spot multispettrale; bande: infrarosso, rosso, verde su rosso, verde, blu; orbita: k62J2597; angolo di ripresa: 4.9 Ovest): zoomata sull'areale del sito di Fabbrica dei Soci di Villabartolomea ripreso da sud (in basso i corsi della Fossa Maestra e del Tartaro; al centro appare, ben visibile, la traccia del tratto orientale dell'argine del sito, cfr. b) (Elaborazione e dati tecnici di C. Giaggio — Regione Veneto — Dipartimento per le Foreste e l'Economia Montana).
b) Foto aerea di Fabbrica dei Soci di Villabartolomea e mappa delle unità operative di base del campo 1 sovrapposta in opaco (in bianco, sul fossato ad ovest del campo, è indicata la linea della sezione stratigrafica di 83 m. Immagine ottenuta con il programma Optilab (cfr. testo): equalizzazione, rimappatura di 16 livelli di grigio su scala cromatica "temperatura dell'acqua", inversione. (Foto aerea C.G.R. — Parma — 13/5036 — scala originale del fotogramma di 1:16600 circa; Conc. S.M.A. n. 1206 del 28/11/89).



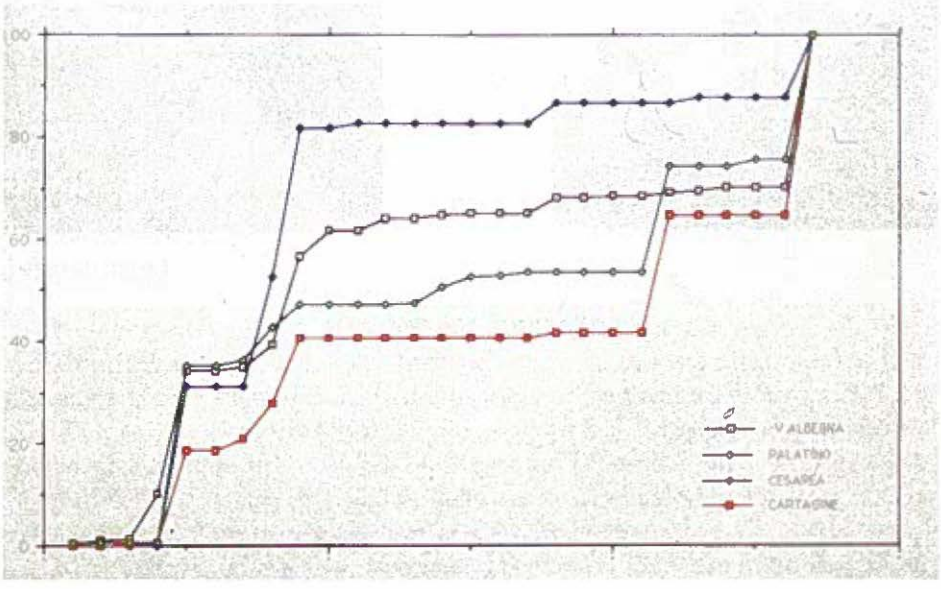


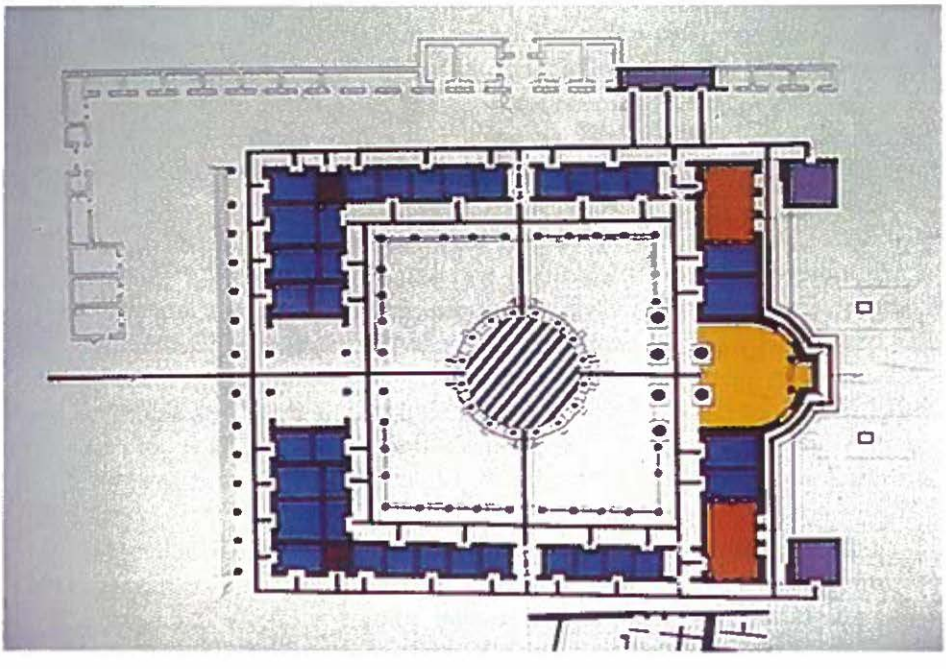
Tav. VIII — a) I quattro siti arginati dell'areale operativo del survey. Le immagini (16 livelli di grigio rimappati su palette "vivid") costituiscono degli "oggetti" grafici (con scipts associati) di una "card" del programma SuperCard (cfr. testo): facendo click su uno dei siti si apre una prima finestra informativa sullo stesso.
b) Immagine digitalizzata relativa al transetto per il survey di "bassa intensità" fra i siti di F. Paviani e Fabbrica dei Soci (da foto aerea I.G.M. conc. S.M.A. 139 del 5/6/1979; scala originale del fotogramma di circa 1:29000).

Tav. IX — a) Programma Digital Darkroom (cfr. testo): immagine digitalizzata dell'areale del sito arginato di Fondo Paviani a 16 livelli di grigio sottoposta ad evidenziazione di soglie (sharpening), equalizzazione, posterizzazione (8 livelli di grigio), tracciatura grossolana (per "features" da 20 pixel minimi di larghezza e lunghezza), incollaggio di immagine tracciata (la "source" scurisce il "target" al 50%), scalatura, incollaggio (default) di scalimetro, manipolazione di luminosità e contrasto (cfr. testo.).
b) Programma Pixel Paint (SuperMac Technology): immagine digitalizzata del sito arginato di Fondo Paviani a 16 livelli di grigio sottoposta a equalizzazione, evidenziazione di soglie (sharpening), tracciatura e rimappatura su palette "metallica". La struttura dell'argine appare ben evidenziata nello zoom di destra.
Fondo Paviani (Legnago-VR): rappresentazione diagrammatica dei rapporti stratigrafici fra due morfotracce contigue al sito arginato. Le tre aree evidenziate a sinistra (in senso antirorario: a) paleo-dosso; b) vincolo di sovrapposizione stratigrafica; c) paleo-argine artificiale) su foto aerea trattata (equalizzazione; "sharpening") vengono riproposte a destra, dopo un ulteriore trattamento di image enhancing (rimappatura su palette cromatica "spettro" e "tracciatura"), in un formato "pseudo-harrisiano" (cfr. testo)
(Da foto aeree C.G.R. — Parma — 12/5070 — scala originale dei fotogramma di 1:16600 circa; Conc. S.M.A. nn. 1205-1206 del 28/11/89).



Tav. Xa-b — Elaborazione a colori computerizzata dei diagrammi cumulativi di p. 116.



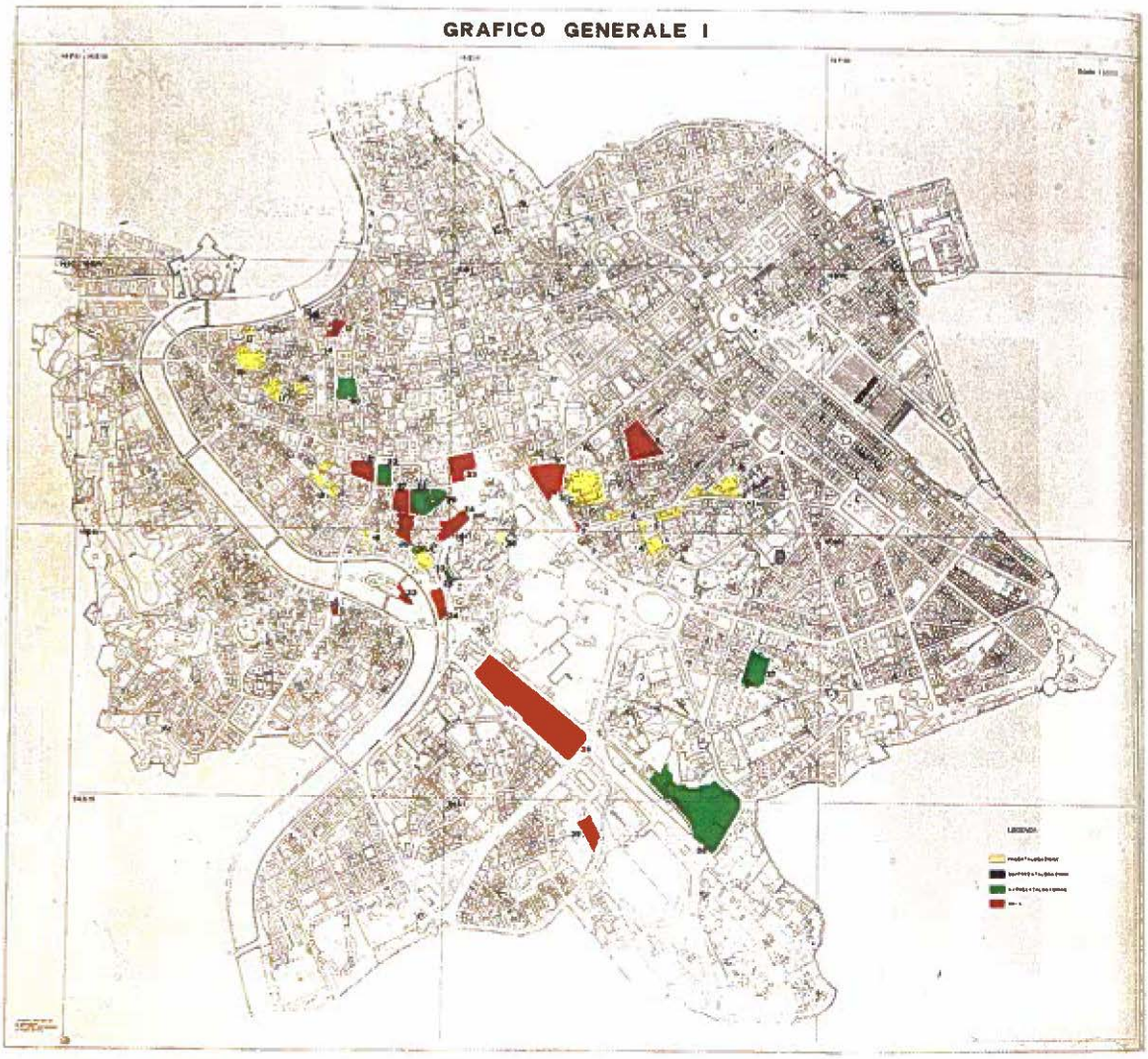


Tav. XIa — Elaborazione a colori computerizzata del diagramma cumulativo di p. 118.

Tav. XIb — Pozzuoli, *Macellum*. Pianta tematica, da un originale in scala 1:200.

a

b



c

Tav. XII — a) Carta delle fortificazioni medioevali tiberine, quadro NE. b) Carta delle fortificazioni medioevali tiberine, quadro SO. c) Quadro di unione degli interventi di schedatura.