

ΜΗΧΑΝΙΚΗ ΟΡΑΣΗ

Χειμερινό Εξάμηνο 2013 – 2014

Εργαστηριακή Άσκηση 2 - «Αναγνώριση Σχημάτων»

Μπαρμπεράκης Στέλιος – 2008030116 – chefarov@gmail.com

1. Εισαγωγή:

Στην παρούσα άσκηση προσπαθήσαμε να αναγνωρίσουμε τις όμοιες εικόνες από ένα σύνολο εικόνων(dataset), βασιζόμενοι στην ανίχνευση σχήματος, χρησιμοποιώντας τις τεχνικές

1. Φάσμα προτύπου (pattern spectrum)
2. Συντελεστές Fourier του διανύσματος του περιγράμματος
3. Συντελεστές Fourier του διανύσματος της παραγώγου της γωνίας σε κάθε σημείο του περιγράμματος.

Το σχήμα είναι από τα πιο σημαντικά χαρακτηριστικά στην ταξινόμηση/ανάκτηση εικόνων βάση του περιεχομένου τους, δηλαδή το συμπερασμό ως προς το τι αναπαριστά μια εικόνα.

2. Οδηγίες Εκτέλεσης – Ροή προγράμματος – Παρουσίαση αποτελεσμάτων:

Το κεντρικό αρχείο της υλοποίησης είναι το αρχείο project2.m. Επειδή οι εικόνες κρατούνται στην μνήμη (για μείωση του χρόνου εκτέλεσης λόγω των πολλών επαναληπτικών πειραμάτων που εκτελέστηκαν) έχουμε ανάγκη για αρκετά MB (200-300MB αρκούν λογικά) διαθέσιμα στη RAM. Τα directories προσπελούνται ως relative paths, και χρησιμοποιήθηκαν Unix paths '/' , οπότε μάλλον πρέπει να αλλάξει αυτό σε '\' για να τρέξει σε Windows.

Η εκτέλεση του script: project2.m περιλαμβάνει τις τρεις μεθόδους σύγκρισης των εικόνων (στις οποίες διοχετεύεται μεταξύ άλλων κι ένα διάνυσμα με τα αναγνωριστικά-ids των εικόνων στόχων – targets).

Έπειτα εκτελείται η σμίκρυνση κατά 50%, επαναλαμβάνεται η παραπάνω διαδικασία. Ακολουθεί περιστροφή 90 μοιρών και ξανα-επαναλαμβάνεται η παραπάνω διαδικασία.

Τα αποτελέσματα βρίσκονται ταξινομημένα στους φακέλους. Όσον αφορά την σύγκριση των τεχνικών ανίχνευσης/ταξινόμησης των εικόνων ελέγχουμε τους φακέλους: spectrum, fourier_coords, fourier_angle, στους οποίους βρίσκεται κι ένα spreadsheet με τον υπολογισμό της απόδοσης για την κάθε μέθοδο.

Η αρίθμηση των εικόνων ακολουθεί την παρακάτω μορφή:

<id_target>.pgm (Η κάθε εικόνα στόχος – κλάση)

<id_target>_<sorting_id>_<id_image>.pgm (Η κάθε μία από τις εικόνες που είχαν την μικρότερη κατά sorting_id απόσταση από το τ στόχο)

id_target: Τα αναγνωριστικά των εικόνων που επιλέχθηκαν ως στόχοι (1,13,27,36,43,54,66,79,89)

sorting_id: Αύξων μετρητής ομοιότητας με το στόχο

id_image: Το αναγνωριστικό της εικόνας που απεικονίζεται σε αυτό το αρχείο.

Άρα η εικόνα 13_1_16.pgm σημαίνει ότι: Η 16η εικόνα είναι η 1η πιο κοντινή σε σχέση με την 13η εικόνα, όπου η 13 είναι ο στόχος. Έπειτα οι εικόνες μιλάνε από μόνες τους...

Επιλέχθηκαν 9 εικόνες στόχοι, για τις οποίες βρίσκουμε τις 9 πιο όμοιες. Πολλές εικόνες έχουν λιγότερες από 10 όμοιες (11 μαζί με την ίδια την εικόνα), οπότε για αυτό κατεβάσαμε το num_similar (αναζήτηση ομοίων) σε 9. Για παράδειγμα υπάρχουν 9 κόκκαλα άρα 8 όμοιες για κάποιο από αυτά.

Μετρική εικόνων και απόδοση μεθόδου (βλ. spreadsheet):

Για κάθε εικόνα-στόχο σχηματίζουμε ένα διάνυσμα 9 θέσεων με μηδενικά και άσσους, όπου άσσο έχουμε στις θέσεις(p) όπου η εικόνα που εμφανίζεται είναι όμοια με τον στόχο και μηδενικό διαφορετικά. Το σκορ ανά εικόνα ορίζεται ως:

$$si = (1.20 * p1 + 1.15 * p2 + .. + 0.85 * p8 + 0.80 * p9) / 9 .$$

Η απόδοση της μεθόδου είναι το άθροισμα των 9 διαφορετικών παραπάνω si δια εννέα, δηλαδή ο μέσος όρος.

Το σκορ ανά εικόνα επιλέχθηκε σταθμισμένο για να δίνεται περισσότερη έμφαση στο αποτέλεσμα που αντιστοιχεί στο 1ο γειτονικό στοιχείο του στόχου σε σχέση με το τελευταίο (9ο). Λεπτομέρεια βέβαια, σε σχέση με την τελική έκβαση του πειράματος...

3. Υλοποίηση μεθόδων

A. Pattern Spectrum:

Σε αυτήν την μέθοδο, η σύγκριση των εικόνων βασίζεται στο άθροισμα των μη-μηδενικών στοιχείων που προκύπτουν μετά την εφαρμογή διαδοχικών opening/closings στις εικόνες του Dataset. Κρίσιμο χαρακτηριστικό επομένως αποτελεί το μέγεθος της εικόνας.

Ανοίγουμε κάθε εικόνα από το Dataset και εφαρμόζουμε διαδοχικά 10 Openings και 10 closings. Φτιάχνουμε ένα διάνυσμα Spectrums διαστάσεων 100×21 όπου κάθε γραμμή αντιστοιχεί στην αντίστοιχη εικόνα με αυτό το αναγνωριστικό (id στο filename) από το dataset. Στις 10 πρώτες στήλες είναι ο αριθμός των μη-μηδενικών pixels που προέκυψαν μετά από κάθε opening, στην επόμενη στήλη ο αριθμός των μη-μηδενικών εικονοστοιχείων της αρχικής εικόνας, και στις επόμενες δέκα ο αντίστοιχος αριθμός μετά από κάθε closing.

Έπειτα κατασκευάζουμε το διάνυσμα distances (9×100) όπου στην θέση i,j αποθηκεύεται η ευκλείδεια απόσταση των διανυσμάτων φάσματος (γραμμών από το spectrums) της εικόνας i και της εικόνας j . Οι γραμμές αντιστοιχούν προφανώς στις εικόνες στόχους

Τέλος ταξινομούμε την κάθε γραμμή σε αύξουσα σειρά και παίρνουμε από την 2η μέχρι την 10η μικρότερη τιμή(στήλη) για κάθε γραμμή, κατασκευάζοντας έτσι ένα διάνυσμα 9×9 (S) όπου βρίσκονται οι εικόνες που επιλέχθηκαν σαν όμοιες (στήλες) για τον κάθε στόχο(γραμμή). Παραλείπουμε την 1η τιμή σε κάθε γραμμή διότι αυτή είναι 0 καθώς αντιστοιχεί στην ίδια την εικόνα στόχο.

B: Μέθοδοι βασιζόμενοι στην ανάλυση Fourier του περιγράμματος

Η ανάλυση σε fourier μας επιτρέπει να συγκρίνουμε τις εικόνες βάση συχνοτικού περιεχομένου. Ουσιαστικά είναι η αναπαράσταση του σχήματος στην συχνότητα. Δύο βασικές κατηγορίες αναπαράστασης σχήματος είναι η αναπαράσταση όλης της περιοχής (region based shape representation), όπου λαμβάνουμε υπόψιν κάθε pixel του αντικειμένου, και η αναπαράσταση περιγράμματος (contour based shape representation). Στην παρούσα εργασία ασχολούμαστε μόνο με την 2η.

Αρχικά υπολογίζουμε το περίγραμμα (κλειστή καμπυλη) της εικόνας

χρησιμοποιώντας την τεχνική `contourc`, κατά την οποία αφαιρούμε από την εικόνα, το αποτέλεσμα του `erosion` της εικόνας με τη μάσκα 3×3 που χρησιμοποιήσαμε και στην προηγούμενη μέθοδο. Ο αντίστοιχος κώδικας βρίσκεται στο αρχείο `boundary.m`

Έπειτα προσπαθήσαμε να ακολουθήσουμε το περίγραμμα στοιχείο προς στοιχείο ώστε να κατασκευάσουμε το μονοδιάστατο διάνυσμα περιγράμματος όπου σε κάθε θέση του βρίσκεται η πληροφορία για τις συντεταγμένες του i -οστού στοιχείου. Επισκεπτόμενοι τον 1ο 8-connected γείτονα που βρίσκαμε κάθε φορά κινούμενοι με καθορισμένη φορά σε όλες τις εικόνες (clockwise), η μέθοδος μας απέτυχε λόγω του ότι το περίγραμμα που μας επέστρεφε η `contourc` δεν ήταν πάχους 1-pixel με αποτέλεσμα να εγκλωβίζεται σε loop καθώς ανίχνευε τριγύρω στοιχεία που είχε ήδη επισκεφτεί. Ο αντίστοιχος κώδικας βρίσκεται στο αρχείο `follow_boundary.m`

Οπότε καταλήξαμε να χρησιμοποιήσουμε την συνάρτηση `bwtraceboundary` της `matlab` (δοκιμάστηκε και η `bwboundaries` με ελαφρώς χειρότερα αποτελέσματα), η οποία δέχεται ως είσοδο:

- ✓ όρισμα 1ο: Την εικόνα ή το περίγραμμα της εικόνας που θα ακολουθήσει (ίδια αποτελέσματα), δηλαδή αντικαθιστά και την `contourc`.
- ✓ Όρισμα 2ο: Ένα αρχικό σημείο εκκίνησης.
- ✓ Όρισμα 3ο: Αρχική κατεύθυνση (πχ 'N': North)
- ✓ Όρισμα 4ο: Μέγιστος αριθμός στοιχείων περιγράμματος (Inf: άπειρα)
- ✓ Όρισμα 5ο: Φορά κίνησης (πχ clockwise)

και παράγει ένα $N \times 2$ πίνακα όπου N ο αριθμός των εικονοστοιχείων του περιγράμματος, άρα 5,1 η τετμημένη και 5,2 η τεταγμένη του 5ου κατά σειρά επίσκεψης εικονοστοιχείου.

Τα εν λόγω διανύσματα αποθηκεύονται στο `cell_matrix: boundaries` και χρησιμοποιούνται και από τις δύο τεχνικές που ακολουθούν.

B.1: Αναπαράσταση καμπύλης με χρήση συντεταγμένων (x,y) :

Η απλούστερη μέθοδος, κατά την οποία κατασκευάζουμε το διάνυσμα $U(n) = x(n) + j*y(n)$, δηλαδή $u(:) = P(:,1) + P(:,2)*1j$; στον κώδικα μας, και στο οποίο εκτελούμε μετασχηματισμό `fourier`, παίρνοντας όσους περισσότερους συντελεστές `fourier` ως αποτέλεσμα φροντίζοντας να μην υπερβαίνουμε το μηκος του μικρότερου περιγράμματος που έχουμε (κάπου στο 270 είναι γι αυτό το dataset).

Έπειτα υπολογίζουμε το μέτρο του κάθε συντελεστή fourier, παραλείπουμε τον 1ο ώστε να πετύχουμε ανεξαρτησία μετατόπισης και διαιρούμε με τον 2ο για να έχουμε ανεξαρτησία κλιμάκωσης.

Κατασκευάζουμε έτσι ένα διάνυσμα Magnitudes κατ' αναλογία του διανύσματος spectrums στην προηγούμενη μέθοδο, και συνεχίζουμε ακριβώς με τον ίδιο τρόπο υπολογίζοντας τον πίνακα με τις ευκλείδεις αποστάσεις, τον οποίο ταξινομούμε σε κάθε γραμμή ώστε να πάρουμε τις “ομοιότερες” εικόνες βάση σχήματος.

Συνήθως σε αυτή τη μεθοδολογία προηγείται ένα φιλτράρισμα με median φίλτρο ώστε να περιορίσουμε το θόρυβο στην εικόνα. Στην προκειμένη περίπτωση επειδή έχουμε συμπαγείς εικόνες χωρίς θόρυβο κάτι τέτοιο δεν είναι ωφέλιμο, αντιθέτως ενδέχεται να καταστρέψει και κάποιες λεπτομέρειες του περιγράμματος, όπως ανιχνεύθηκε και πειραματικά.

Επίσης έπεται μια κατωφλίωση στο διάνυσμα των συντελεστών καθώς όσο προχωράμε προς υψηλότερους συντελεστές, κωδικοποιούνται μεγαλύτερες συχνότητες άρα περισσότερο λεπτομέρειες και από ένα σημείο και μετά είναι κυρίως θόρυβος. Στην περίπτωση μας πάλι δεν χρειάζεται κάτι τέτοιο για τον ίδιο λόγο με το φιλτράρισμα. Δοκιμάστηκε πειραματικά και δεν είχαμε αλλαγή στα αποτελέσματα.

B.2 Αναπαράσταση καμπύλης με χρήση γωνιών:

Σε αυτήν την τεχνική χρησιμοποιούμε την πληροφορία που παίρνουμε από το περίγραμμα για την τιμή της γωνίας σε κάθε σημείο του περιγράμματος και ειδικότερα την παράγωγο της γωνίας ώστε να έχουμε ανεξαρτησία περιστροφής.

Διατρέχουμε τον πίνακα με τις συντεταγμένες του περιγράμματος κάθε εικόνας και υπολογίζουμε την γωνία στο κάθε σημείο συγκρίνοντας $x, x+1$ και $y, y+1$. Χρησιμοποιούμε θετικά ακέραια πολλαπλάσια του $\pi/4$.

Υπολογίζουμε την παράγωγο της γωνίας στο κάθε σημείο, δηλαδή το $\Delta\theta = \theta(n+1) - \theta(n)$.

Έπειτα εκτελούμε τις παρακάτω πράξεις πινάκων:

```
dth = dth - 2*pi*(dth>pi);  
dth = dth + 2*pi*(dth<-pi);
```

έτσι ώστε στα σημεία που η διαφορά γωνίας είναι μεγαλύτερη από π ή

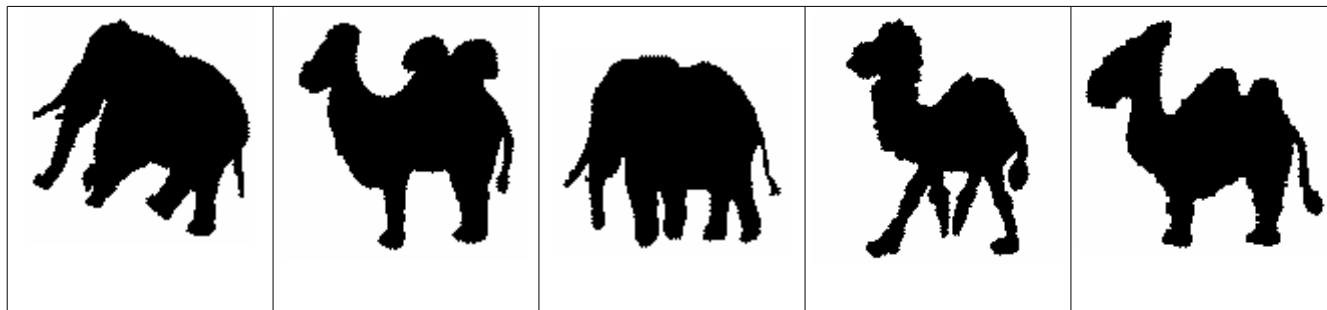
μικρότερη από $-\pi$ να την κανονικοποιήσουμε ως σχετική γωνία με την προηγούμενη κατεύθυνση-γωνία. Έτσι η διαφορά του 'E' (0) με το 'SE' ($7\pi/4$) να γίνει από $7\pi/4 \rightarrow -\pi/4$.

Έπειτα υπολογίζουμε το μετασχηματισμό φουριερ του διανύσματος $\Delta\theta(d\theta)$ για κάθε εικόνα και παίρνουμε τα μέτρα των συντελεστών φουριερ κατασκευάζοντας τον πίνακα magnitudes και επαναλαμβάνοντας την γνωστή πλέον διαδικασία.

4. Σχολιασμός Αποτελεσμάτων – Σύγκριση μεθόδων

A. Pattern Spectrum:

Χαρακτηριστικό της μεθόδου είναι ότι μπερδεύει σχήματα με παρόμοιο μέγεθος, αφού η μόνη μετρική της είναι το άθροισμα των μη-μηδενικών εικονοστοιχείων. Αυτό είναι εμφανές παρατηρώντας τα αποτελέσματα στο φάκελο spectrum όπου φαίνεται ότι η μέθοδος μπερδεύει τα εργαλεία μεταξύ τους και με τα κόκκαλα, τους ελέφαντες με τις καμήλες κλπ.



Παρατηρήσαμε το προβλεπόμενο από την θεωρία γεγονός ότι μετά το 2o Opening ή closing ο αριθμός των εικονοστοιχείων σταθεροποιούνται καθώς το κάθε opening ή closing περιλαμβάνει και dilation και erosion άρα τα διαδοχικά αποτελέσματα αλληλοαναιρούνται.

Δοκιμάσαμε δύο διαφορετικά δομικά στοιχεία (τετράγωνο και κύκλο), και τα αποτελέσματα ήταν να μην προφανώς διαφορετικά αλλά οι διαφοροποιήσεις ήταν σχετικά μικρές με αποτέλεσμα να μην επηρεάζεται σε μεγάλο βαθμό η απόδοση. Στο set εικόνων που δεν υπέστη scale είχαμε ένα 3-4% καλύτερο αποτέλεσμα με το τετράγωνο.

Το ποσοστό απόδοσης 52%, είναι αρκετά ικανοποιητικό για την δεδομένη μέθοδο, και φυσικά εξαρτάται σε τεράστιο βαθμό από το dataset.

Τέλος παρατηρήσαμε ότι η μέθοδος δεν επηρεάζεται από περιστροφή (άρα και μετατόπιση), τουναντίον επηρεάζεται από την κλιμάκωση. Τα παραπάνω είναι αναμενόμενα καθώς με την κλιμάκωση μειώνουμε τον αριθμό των εικονοστοιχείων άρα και των μη-μηδενικών εικονοστοιχείων της εικόνας, ενώ με την περιστροφή όχι.

Μειονεκτήματα μεθόδου:

- Σύγχυση μεταξύ εικόνων με διαφορετικό σχήμα αλλά παρόμοιο μέγεθος.
- Χρονική πολυπλοκότητα λόγω της εφαρμογής των τελεστών erosion και dilation που διατρέχουν κάθε εικονοστοιχείο, επαναληπτικά σε κάθε εικόνα. Περίπου 10 φορές πιο αργή από τις μεθόδους με την ανάλυση fourier στην υλοποίησή μας. Θα μπορούσε να μειωθεί, μειώνοντας τον αριθμό των iterations για το closing Και το Opening μια και όπως είδαμε το φάσμα σταθεροποιείται μετά το 2ο, 3ο opening/closing.
- Μεγάλη εξάρτηση από το dataset, σε ορισμένες περιπτώσεις δεν θα δουλεύει καθόλου καλά.
- Ανάγκη για διαρκή πειραματισμό με τα δομικά στοιχεία κατά την εφαρμογή των τελεστών για κάθε διαφορετικό dataset.

Πλεονεκτήματα μεθόδου:

- Καλή και σταθερή απόδοση για εικόνες όπου το μέγεθος καθορίζει σε μεγάλο βαθμό την οντότητα που αναπαριστάται. Για παράδειγμα ταξινόμηση μηχανοκίνητων (μοτοσυκλέτα, κούρσα, φορτηγό, επαγ. Φορτηγό, κλπ).
- Απλή και γρήγορη υλοποίηση
- Εύκολη αποσφαλμάτωση, που περιλαμβάνει και την απεικόνιση της κάθε εικόνας μετά το opening/closing, σε σύγκριση με απεικόνιση διανυσμάτων κλπ

B.1: Αναπαράσταση καμπύλης με χρήση συντεταγμένων (x,y) :

Σαν μέθοδος εκμεταλλεύεται την σχετική θέση του κάθε σημείου του περιγράμματος σε σχέση με το αρχικό σημείο του περιγράμματος

Η κλιμάκωση δεν αλλάζει το αποτέλεσμα καθώς έχουμε διαιρέσει όλους τους συντελεστές με τον α_1 , η οποία κανονικοποίηση κάνει την παράσταση ανεξάρτητη κλιμάκωσης. Η μετατόπιση επίσης δεν επηρεάζει το αποτέλεσμα καθώς έχουμε αφαιρέσει τον α_0 που περιέχει τη dc συνιστώσα.

Αυτό που θα περιμέναμε είναι η περιστροφή να επηρεάζει το αποτέλεσμα καθώς για να γίνει αυτή η μέθοδος ανεξάρτητη περιστροφής χρειάζεται να βρίσκουμε κάθε φορά τους 2 συντελεστές με το μεγαλύτερο μέτρο και να θέτουμε τη γωνία φάσης τους ίση με το μηδέν (Keyes and Winstanley 1999 σελ.4 - επισυνάπτεται). Αντιθέτως τα αποτελέσματα είναι αναλλοίωτα με την περιστροφή των εικόνων στόχων για κάποιο λόγο.

Η απόδοση της μεθόδου 31.6% είναι σχετικά απογοητευτική καθώς θα περιμέναμε να υπερβαίνει έστω οριακά αυτήν της pattern spectrum, καθώς είναι γνωστό ότι οι Fds αποτελούν κυρίαρχη μεθοδολογία για contour based shape representation.

Μειονεκτήματα μεθόδου:

- Απαίτηση ακριβούς καταγραφής του περιγράμματος.
- Σύγχυση μεταξύ εικόνων με παρόμοιο περίγραμμα αλλά διαφορετικού μεγέθους και διαφορετικού εν τέλει σχήματος.
- Δυσκολία στην αποσφαλμάτωση

Πλεονεκτήματα μεθόδου:

- Χαμηλή χρονική πολυπλοκότητα.
- Καθολική εφαρμογή, χωρίς ανάγκη για ιδιαίτερες μετατροπές ανά dataset
- Έμφαση στο περίγραμμα, καλύτερη αναπαράσταση σχήματος
- Επιτυχής ταξινόμηση εικόνων συγκεκριμένης αναπαράστασης/σχήματος λόγω ανίχνευσης της λεπτομέρειας.

B.2 Αναπαράσταση καμπύλης με χρήση γωνιών:

Όπως αναμενόταν τα αποτελέσματα είναι ανεξάρτητα κλιμάκωσης, μετατόπισης, περιστροφής, λόγω της κανονικοποίησης με α_1 , μηδενισμού του α_0 και υπολογισμό της $\Delta\theta$ (διαφοράς γωνιών) αντίστοιχα.

Η απόδοση 26.79% είναι άκρως απογοητευτική καθώς θα αναμέναμε η εν λόγω μέθοδος να έχει τα μεγαλύτερα ποσοστά επιτυχίας, ή τουλάχιστον καλύτερα από την pattern spectrum.

Πιθανότατα να παραλείψαμε κάποιο βήμα που χρειαζόταν αν και ακολουθήσαμε γενικά πιστά τις οδηγίες.

Επίσης ένα αξιοπερίεργο σημείο με αυτή τη μέθοδο ήταν ότι η αύξηση των συντελεστών fourier από 200 σε 256 έριξε αρκετά(σχεδόν στο μισό) την απόδοση της μεθόδου, σε αντίθεση με την προηγούμενη που την βελτίωσε ανεπαίσθητα (πρακτικά ίδια αποτελέσματα) κάτι το οποίο είναι πιο λογικό.

Μειονεκτήματα μεθόδου:

- Πάνω κάτω τα ίδια με την B.1
- Σχετικά μεγαλύτερη χρονική πολυπλοκότητα από την B.1 λόγω των υπολογισμών των γωνιών και της διαφοράς τους. Ωστόσο παραμένει αρκετά γρήγορη τεχνική σε σύγκριση με άλλες, ειδικά το pattern spectrum...

Πλεονεκτήματα μεθόδου:

- Πάνω κάτω τα ίδια με την B.1
- Εκμεταλλεύεται την κλίση στο κάθε σημείο παρά την σχετική απόσταση από το αρχικό, με αποτέλεσμα να έχουμε περισσότερη έμφαση στις γωνίες και την λεπτομέρεια του περιγράμματος άρα και του σχήματος.