

**ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ  
ΚΕΝΤΡΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ  
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ  
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ**

# **ΨΗΦΙΑΚΗ ΕΠΕΞΕΡΓΑΣΙΑ ΕΙΚΟΝΑΣ**

## **Ενότητα 1.**

Χαράλαμπος Π. Στρουθόπουλος  
Καθηγητής



Ιούλιος 2015

## ΠΕΡΙΕΧΟΜΕΝΑ

1. Κεφάλαιο 1 – Εισαγωγή και ορισμοί
  - 1.1 Είδη ψηφιακών εικόνων.
  - 1.2 Βασικές έννοιες
  - 1.3 Ιστόγραμμα

# ΚΕΦΑΛΑΙΟ 1

## Εισαγωγή και ορισμοί

Η ψηφιακή επεξεργασία εικόνας (ΨΕΕ) αποτελεί έναν ευρύ επιστημονικό κλάδο που αναπτύχθηκε με την ραγδαία εξέλιξη των υπολογιστών. Ο όρος εικόνα χρησιμοποιείται ευρύτερα από την απλή απεικόνιση ενός σκηνικού και περιλαμβάνει την αποτύπωση κάθε είδους πληροφοριών. Τα υπερηχογραφήματα, οι μαγνητικές τομογραφίες, οι δορυφορικές φωτογραφίες κ.α. μπορούν να επεξεργαστούν ως ψηφιακές εικόνες.

Οι στόχοι της ΨΕΕ είναι οι εξής:

- 1) Η ψηφιοποίηση και κωδικοποίηση εικόνων με σκοπό την αποθήκευση, μετάδοση και εκτύπωσή τους.
- 2) Η βελτίωση και η αποκατάσταση των εικόνων με σκοπό την καλύτερη απεικόνισή τους.
- 3) Η ανάλυση και κατανόηση των εικόνων

Η ΨΕΕ συνεργάζεται με τους παρακάτω επιστημονικούς κλάδους:

- 1) Ψηφιακή Επεξεργασία Σημάτων (ΨΕΣ)
- 2) Ρομποτική όραση
- 3) Τεχνητή Νοημοσύνη
- 4) Αναγνώριση Προτύπων
- 5) Νευρωνικά Δίκτυα
- 6) Ασαφής Λογική
- 7) Κωδικοποίηση
- 8) Γραφικά Η/Υ

Η ψηφιακή εικόνα είναι ένα πεπερασμένο σύνολο περιοχών όπου κάθε περιοχή είναι χρωματισμένη με χρώμα που προέρχεται από ένα πεπερασμένο σύνολο χρωμάτων. Στις περισσότερες των περιπτώσεων, μια ψηφιακή εικόνα είναι ένα ορθογώνιο, διαιρεμένο με γραμμές και στήλες σε ορθογώνιες περιοχές που κάθε μία έχει συγκεκριμένο χρώμα. Μια τέτοια περιοχή ονομάζεται στοιχείο της εικόνας ή εικονοστοιχείο. Στην αγγλική λέγεται *pixel* ή *pel*, όρος ο οποίος προέρχεται από τη σύντμηση των λέξεων *picture element*. Από τα παραπάνω καταλαβαίνουμε πως μια ψηφιακή εικόνα είναι ένα διδιάστατο ψηφιακό σήμα.

Αν κάθε χρώμα κωδικοποιηθεί με έναν αριθμό τότε η ορθογώνια ψηφιακή εικόνα περιγράφεται από έναν πίνακα αριθμών  $J \times K$ , όπου  $J$  το πλήθος των γραμμών και  $K$  το πλήθος των στηλών της ψηφιακής εικόνας. Η τιμή  $I(j,k)$  με  $k=0,1,2,\dots,K-1$  και  $j=0,1,2,\dots,J-1$  είναι ο κωδικός του χρώματος της ψηφιακής εικόνας.

Η μετατροπή μιας εικόνας σε ψηφιακή μορφή ουσιαστικά είναι η μετατροπή ενός δισδιάστατου αναλογικού σήματος σε ψηφιακό και απαιτεί τις διαδικασίες της δειγματοληψίας και του κβαντισμού.

### 1.1 Είδη ψηφιακών εικόνων

Υπάρχουν τρία είδη ψηφιακών εικόνων που χαρακτηρίζονται από το πλήθος των χρωμάτων που περιέχουν:

- 1) *Δυαδικές εικόνες (binary images)*: Κάθε εικονοστοιχείο των εικόνων μπορεί να χρωματιστεί με ένα από δύο χρώματα (συνήθως άσπρο ή μαύρο). Για κάθε εικονοστοιχείο απαιτείται ένα bit πληροφορίας, π.χ. με τιμή μηδέν (0) για το μαύρο και ένα (1) για λευκό. Οι εικόνες των εγγράφων που αποτελούνται μόνο από το χρώμα του χαρτιού και της μελάνης αναπαρίστανται σε δυαδική ψηφιακή μορφή.
- 2) *Εικόνες αποχρώσεων του γκρι (gray level images)*: Κάθε εικονοστοιχείο των εικόνων μπορεί να χρωματιστεί με μία από τις αποχρώσεις του γκρι οι οποίες ξεκινούν από το μαύρο και καταλήγουν στο λευκό. Από αυτές τις αποχρώσεις συνήθως λαμβάνονται 256 αντιπροσωπευτικές που κωδικοποιούνται με τιμές  $0,1,2,\dots,255$ . Η απόχρωση κάθε εικονοστοιχείου προφανώς απαιτεί πληροφορία ενός byte.
- 3) *Έγχρωμες εικόνες (color images)* στις οποίες κάθε εικονοστοιχείο χρωματίζεται με χρώματα που προέρχονται από την ανάμειξη των αποχρώσεων του κόκκινου, πράσινου και μπλε (*RGB*). Για κάθε ένα από τα τρία αυτά χρώματα λαμβάνονται 256

αποχρώσεις δηλαδή πληροφορία του ενός byte. Συνεπώς κάθε εικονοστοιχείο της έγχρωμης εικόνας, απαιτεί 3 bytes.

## 1.2 Βασικές έννοιες

Το σύνολο των χρωμάτων που μπορούν να χρησιμοποιηθούν για τον χρωματισμό των εικονοστοιχείων της εικόνας λέγεται χρωματική παλέτα. Εάν  $C$  είναι το πλήθος των χρωμάτων, τότε για την κωδικοποίησή τους απαιτούνται  $B$  bits και ισχύουν οι σχέσεις

$$C=2^B \Leftrightarrow B=\log_2 C$$

Το  $B$  ονομάζεται *βάθος bit (bit depth)* της ψηφιακής εικόνας. Εάν η εικόνα έχει  $K$  στήλες και  $J$  γραμμές τότε για την απεικόνισή της απαιτούνται  $J \times K \times B$  bits. Ο Πιν.1.2.1 παρουσιάζει ενδεικτικές τιμές των παραπάνω μεγεθών.

Είδος εικόνας	J	K	B	Bits	bytes
Διαδική	100	100	2	20000	2500
Αποχρώσεων του γκρι	100	100	8	80000	10000
Έγχρωμη RGB	100	100	24	240000	30000
Διαδική	512	512	2	524288	65536
Αποχρώσεων του γκρι	512	512	8	2097152	262144
Έγχρωμη RGB	512	512	24	6291456	786432

Πίνακας 1.2.1.

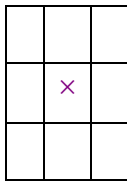
*Ευκρίνεια της εικόνας* είναι το πλήθος των εικονοστοιχείων ανά μονάδα επιφάνειας και καθορίζει πόσο λεπτομερής είναι η ψηφιακή αναπαράσταση της εικόνας. Η ευκρίνεια  $E$  μιας εικόνας διαστάσεων  $J \times K$  και εμβαδού  $A$  δίνεται από την σχέση

$$E = \frac{J \times K}{A} \quad (1.2.1)$$

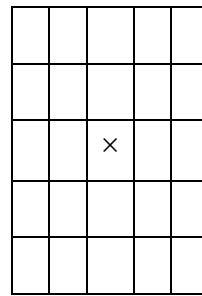
και μετρείται σε πλήθος εικονοστοιχείων ανά μονάδα επιφάνειας π.χ. pixels/mm<sup>2</sup> ή dpi ( dots per inch : κουκίδες ανά ίντσα). Η ευκρίνεια εξαρτάται τόσο από το πλήθος των εικονοστοιχείων όσο και από τις φυσικές διαστάσεις της εικόνας.

Μια ομάδα γειτονικών εικονοστοιχείων λέγεται γειτονιά. Σε μια γειτονιά  $S_{M \times N}$  με  $M$  γραμμές και  $N$  στήλες μιας εικόνας διαστάσεων  $J \times K$  υπάρχει ένα κεντρικό εικονοστοιχείο  $(j_c, k_c)$  όταν  $M, N$  είναι περιττοί αριθμοί. Η θέση των εικονοστοιχείων της  $S$  αναφέρονται συχνά, σχετικά με την θέση του κεντρικού εικονοστοιχείου της. Η πιο συνήθης γειτονιά

είναι τριών (3) γραμμών και τριών (3) στηλών και λέγεται γειτονιά  $3 \times 3$ . Στο Σχ.1.2.1 φαίνονται  $3 \times 3$  και  $5 \times 5$  γειτονιές.



Γειτονιά 3×3



Γειτονιά 5×5

Σχήμα 1.2.1: Γειτονιές 3×3 και 5×5 με τα κεντρικά τους εικονοστοιχεία.

Σε κάθε εικονοστοιχείο μιας γειτονιάς με  $M$  γραμμές και  $N$  στήλες όπου  $M$  και  $N$  περιττοί ακέραιοι, μπορούμε να αντιστοιχίσουμε έναν συντελεστή  $w(m,n)$ ,  $m=-(M-1)/2, \dots, (M-1)/2$ ,  $n=-(N-1)/2, \dots, (N-1)/2$  στην γραμμή

$$r = (M-1)/2 + m + 1, r = 1 \dots M,$$

και στην στήλη

$$c = (N-1)/2 + n + 1, c = 1 \dots N,$$

Ονομάζουμε μάσκα  $W$  τον πίνακα  $S$  με στοιχεία τις τιμές τις μάσκας σύμφωνα με την σχέση  $s_{rc} = w(m,n)$ .

Για παράδειγμα για  $N=M=3$  ο πίνακας  $S$  είναι

$$\begin{vmatrix} w(-1,-1) & w(-1,0) & w(-1,1) \\ w(0,-1) & w(0,0) & w(0,1) \\ w(1,-1) & w(1,0) & w(1,1) \end{vmatrix}$$

Οι συντελεστές της μάσκας και οι τιμές φωτεινότητας των εικονοστοιχείων μπορούν να εμπλακούν σε χρήσιμους υπολογισμούς για την επεξεργασία της εικόνας. Ο συνηθέστερος υπολογισμός δίνεται από την σχέση

$$A(j,k) = \sum_m \sum_n w(m,n) \cdot I(j+m, k+n)$$

όπου  $(j,k)$  εικονοστοιχείο της εικόνας  $I$ .

Στο παρακάτω σχήμα δίνονται μάσκες 3×3 και η εφαρμογή τους σε μία 3×3 περιοχή εικονοστοιχείων.

$$1/9 \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \quad \begin{array}{|c|c|c|} \hline 240 & 250 & 240 \\ \hline 240 & 100 & 150 \\ \hline 240 & 250 & 250 \\ \hline \end{array}$$

$$\frac{240 + 250 + 240 + 240 + 100 + 150 + 240 + 250 + 250}{9} \approx 191$$

$$1/17.64 \begin{array}{|c|c|c|} \hline 1.41 & 1 & 1.41 \\ \hline 1 & 8 & 1 \\ \hline 1.41 & 1 & 1.41 \\ \hline \end{array} \quad \begin{array}{|c|c|c|} \hline 240 & 250 & 240 \\ \hline 240 & 100 & 150 \\ \hline 240 & 250 & 250 \\ \hline \end{array}$$

$$\frac{1.41 \cdot 240 + 250 + 1.41 \cdot 240 + 240 + 8 \cdot 100 + 150 + 1.41 \cdot 240 + 250 + 1.41 \cdot 250}{17.64} \approx 138$$

Σχήμα 1.2.2.

Αν  $(j_c, k_c)$  είναι το κεντρικό εικονοστοιχείο της γειτονιάς  $S$  της εικόνας  $I_{J \times K}$  η τιμή  $A$  μπορεί να αποδοθεί ως τιμή φωτεινότητας του εικονοστοιχείου  $(j_c, k_c)$  μιας νέας εικόνας  $I'_{J \times K}$ . Αν αυτό εφαρμοσθεί για όλες τις γειτονιές της εικόνας  $I$ , τότε λέμε ότι η νέα εικόνα  $I'$  προέκυψε από το φιλτράρισμα της  $I$  με την μάσκα  $W$ . Αν περιγράψουμε μια τέτοια πράξη με όρους της ψηφιακής επεξεργασίας σήματος τότε η εφαρμογή της μάσκας  $W$  σε όλα τα εικονοστοιχεία της εικόνας ισοδυναμεί με την έξοδο ενός γραμμικού και ανεξάρτητου από την μετατόπιση (LTI) συστήματος με απόκριση κρουστικής διέγερσης της μορφής

$$h(j, k) = \sum_m \sum_n w(m, n) \cdot \delta(j + m, k + n)$$

	w(1,1)	w(1,0)	w(1,-1)
<b>h=</b>	w(0,1)	w(0,0)	w(0,-1)
	w(-1,1)	w(-1,0)	w(-1,-1)

και θα δίνεται από τη σχέση  $I' = I * h$

### 1.3. Ιστόγραμμα

Αν η ψηφιακή εικόνα αποτελείται από αποχρώσεις του γκρι μπορούμε να δημιουργήσουμε την κατανομή του πλήθους των εικονοστοιχείων που έχουν την ίδια τιμή απόχρωσης για κάθε



απόχρωση. Η κατανομή αυτή λέγεται ιστόγραμμα των αποχρώσεων της εικόνας και δίνεται μαθηματικά από την σχέση

$$h(g) = \sum_{\forall I(j,k)=g} 1 \quad (1.3.1)$$

όπου  $g=0,1,2..G-1$ ,  $G$  το πλήθος των αποχρώσεων,  $j=0, \dots, J-1$ ,  $k=0, \dots, K-1$ ,  $J$  το πλήθος των γραμμών,  $K$  το πλήθος των στηλών της εικόνας,  $I(j,k)$  η τιμή της φωτεινότητας στο σημείο  $(j,k)$  και  $h(g)$  το πλήθος των εικονοστοιχείων με απόχρωση  $g$ . Το ιστόγραμμα δηλαδή μας πληροφορεί πόσα εικονοστοιχεία υπάρχουν με συγκεκριμένη τιμή απόχρωσης. Η συνάρτηση

$$H(g) = \frac{h(g)}{\sum_{g=0}^{G-1} h(g)} \quad (1.3.2)$$

λέγεται κανονικοποιημένο ιστόγραμμα και δίνει την πιθανότητα ένα τυχαίο εικονοστοιχείο της εικόνας να έχει απόχρωση  $g$ .