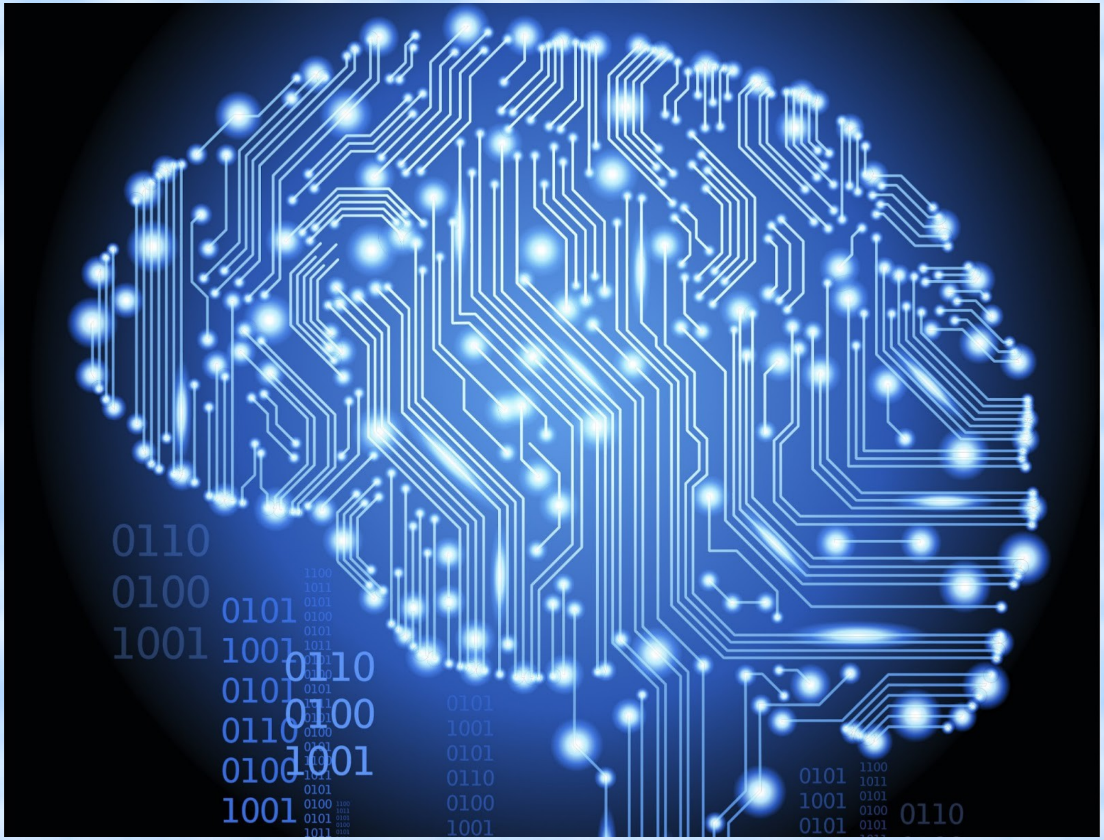
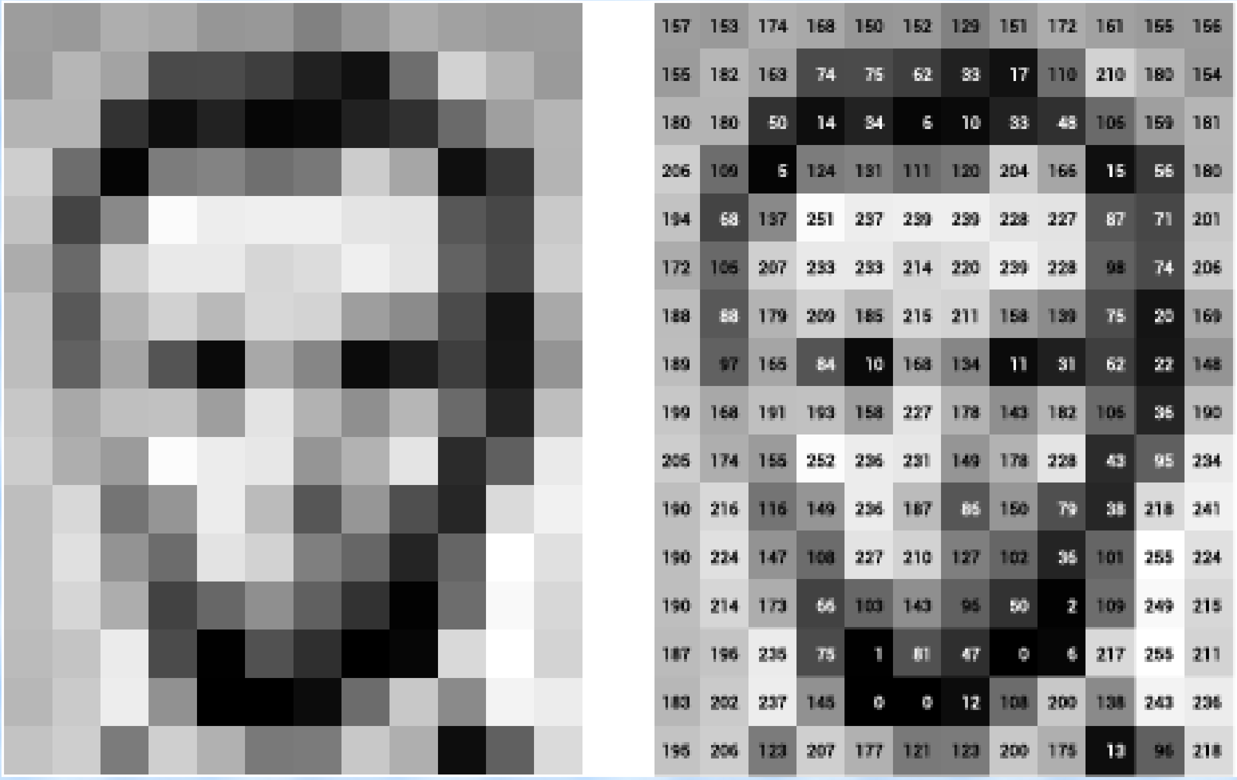
Α.Τ.Ε.Ι. ΜΑΚΕΔΟΝΙΑΣ
Τμήμα πληροφορικής και επικοινωνιών

Συμπίεση ψηφιακών εικόνων με ανάλυση κύριων συνιστωσών και χρήση νευρωνικού δικτύου.



Ψηφιακή είκόνα

Η ψηφιακή εικόνα είναι ένα πεπερασμένο σύνολο περιοχών όπου κάθε περιοχή είναι χρωματισμένη με χρώμα που προέρχεται από ένα πεπερασμένο σύνολο χρωμάτων.

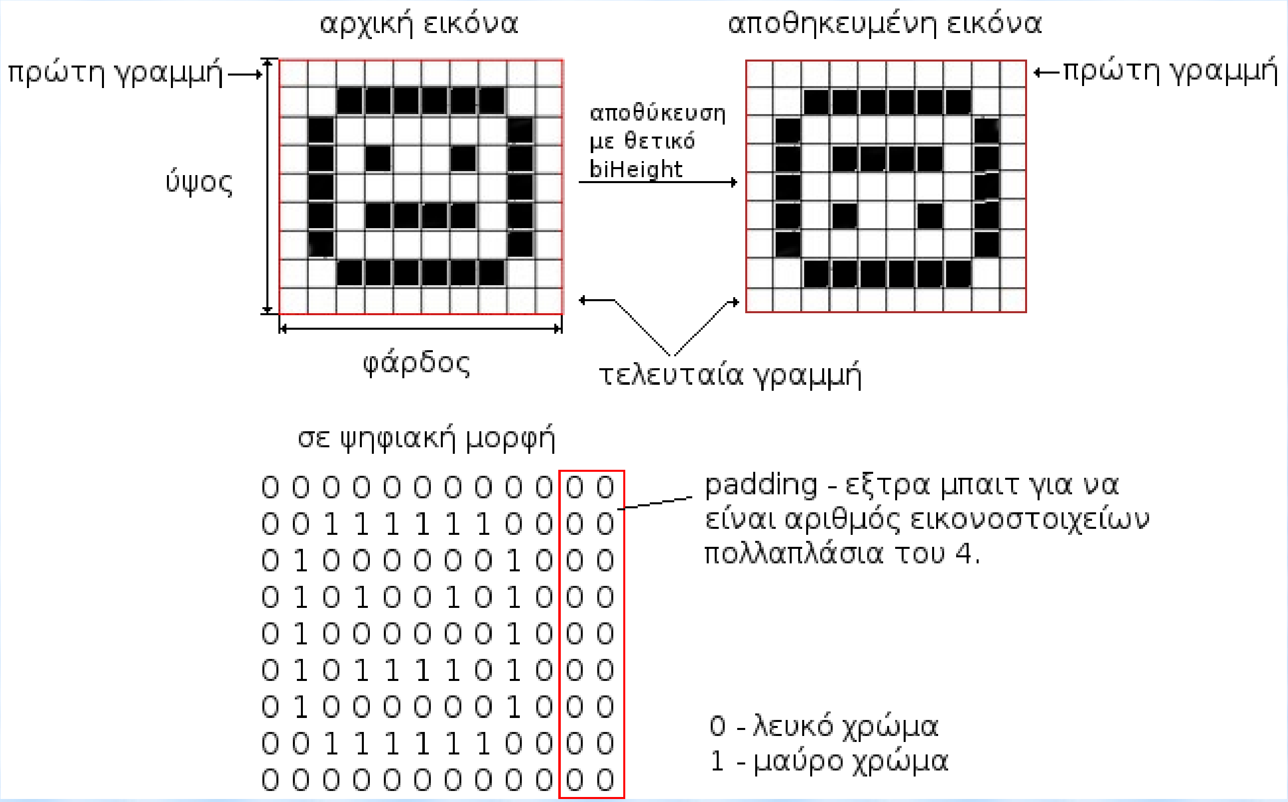


157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

Ψηφιακή εικόνα

Μια τέτοια περιοχή ονομάζεται στοιχείο της εικόνας ή εικονοστοιχείο(pixel ή pel από το picture element).

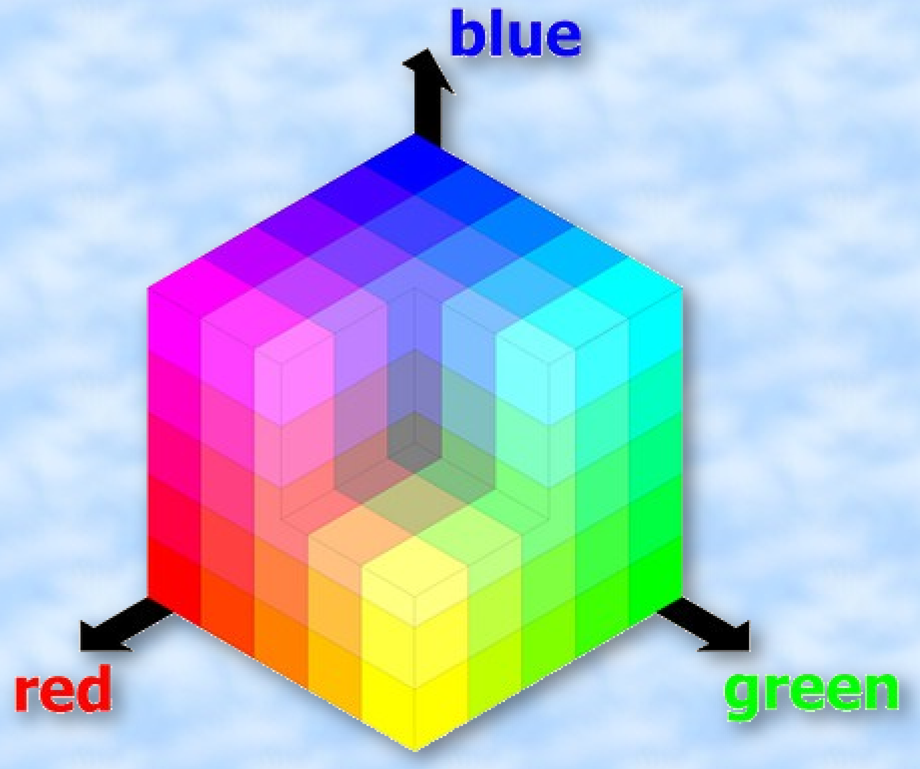
Πίνακας εικονοστοιχείων της εικόνας bitmap.



Βάθος χρώματος

Παλέτα χρωμάτων

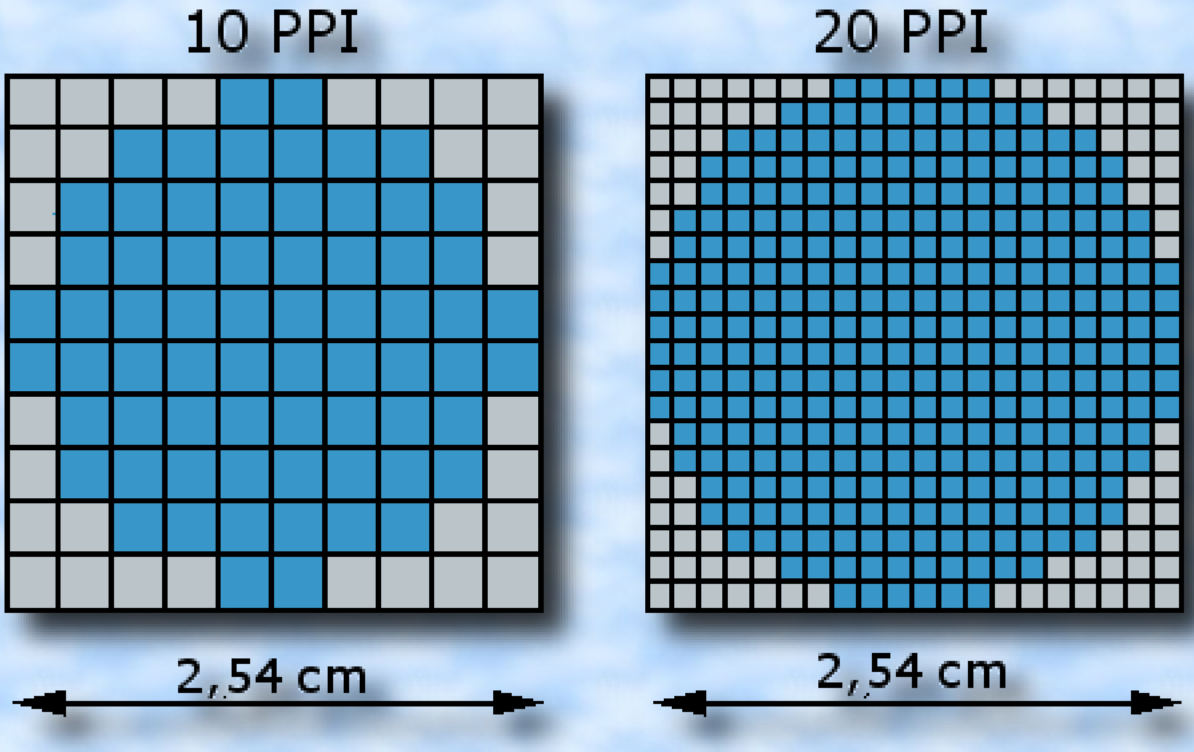
Βάθος χώματος 2- 48 bit
2 bit – 2 χρώματα
8 bit - "ασπρόμαυρη"
16 bit – 65536 χρώματα
24 bit – true color
32 bit – 24bit + alpha channel
48 bit – image processing



Ευκρίνεια

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×K και εμβαδού A δίνεται από την σχέση:

$$E = \frac{J \times K}{A}$$



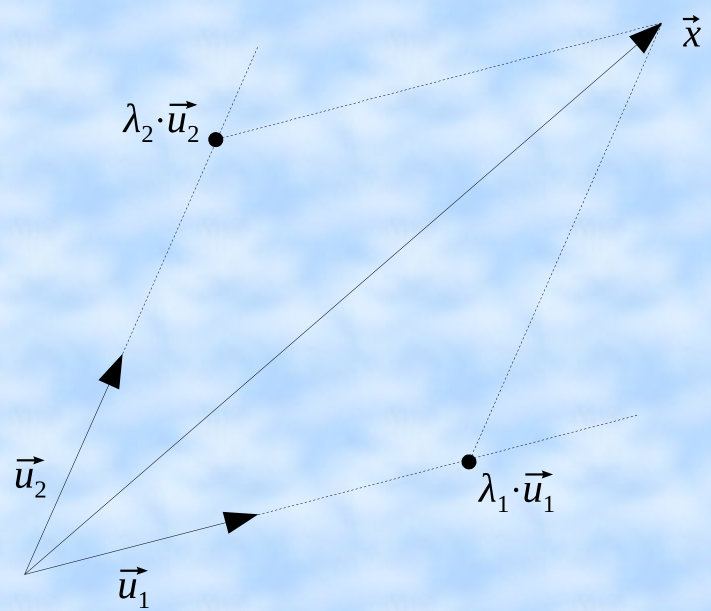
Συμπίεση

- LZW – Lambel-Ziv-Welch
- RLE – Run Length encodign
- Deflation - LZ77+Huffman
- Transform coding
- Fractal
- Color space reducing

PCA

Με τον όρο ανάλυση εννοούμε συνήθως την εύρεση των συστατικών ενός πράγματος ή την ύπαρξη συστατικών μέσα σ' αυτό. Παρόμοια χρήση έχει ο όρος αυτός και στα μαθηματικά. Έτσι ξέρουμε, ότι τα διανύσματα αναλύονται σε άλλα διανύσματα με τον κανόνα του παραλληλογράμμου π.χ.

$\vec{x} = \lambda_1 \cdot \vec{u_1} + \lambda_2 \cdot \vec{u_2}, \mathit{οπου}\ \lambda_{1,}\ \lambda_2 \in \mathbb{R}$



Δεδομένων των $\vec{u_1}, \vec{u_2}$ ζητούνται τα λ_1, λ_2.

PCA

Η παραπάνω διαδικασία απλοποιείται σημαντικά, αν τα $\vec{u}_1, \vec{u}_2$ και γενικότερα $\vec{u}_\kappa, \kappa = 1,2,3 \ldots K$ είναι κάθετα μεταξύ τους και έχουνε μέτρο(μήκος) την μονάδα (αποτελούν ορθοκανονική βάση). Η εύρεση των $\lambda_1, \lambda_2, \lambda_3 \ldots \lambda_\kappa$ τότε είναι απλά υπολογισμός του εσωτερικού γινομένου $\lambda_\kappa = \langle \vec{x}, \vec{u}_\kappa \rangle$

$$\lambda_1 = |\vec{\alpha}||\vec{u}_1| \cdot \cos\theta = \langle \vec{\alpha}, \vec{u}_1 \rangle$$

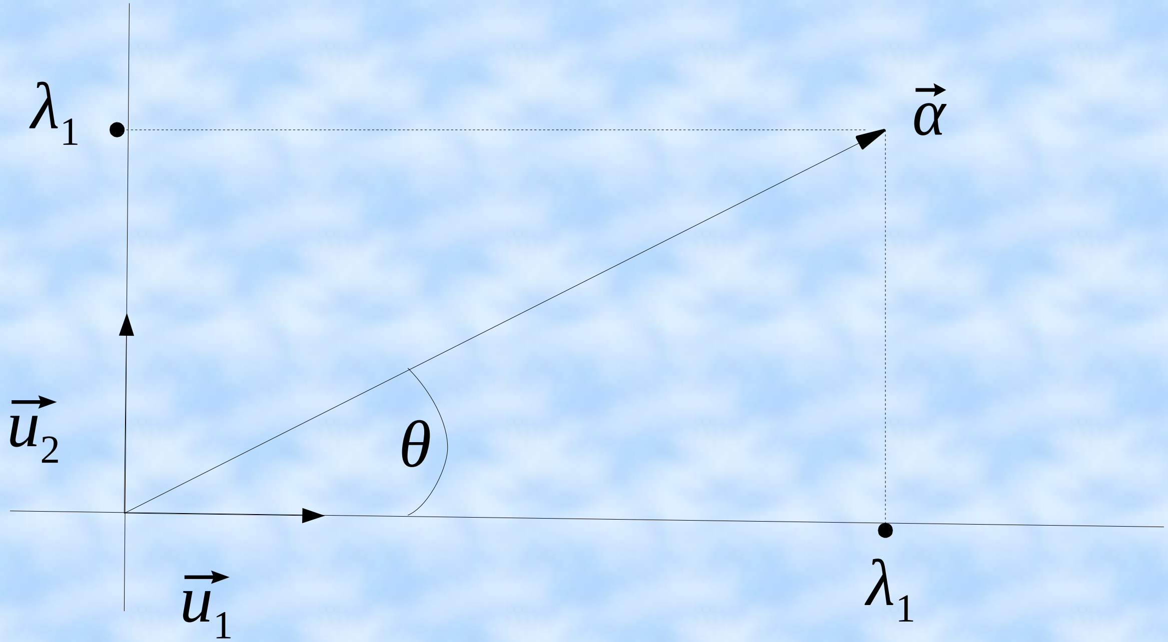
$$\lambda_2 = |\vec{\alpha}||\vec{u}_1| \cdot \cos\theta = \langle \vec{\alpha}, \vec{u}_2 \rangle$$

ή αλγεβρικά για

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}, u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, u_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\langle \alpha, u_1 \rangle = \alpha_1 \cdot 1 + \alpha_2 \cdot 0 = \alpha_1 = \lambda_1$$

$$\langle \alpha, u_2 \rangle = \alpha_1 \cdot 0 + \alpha_2 \cdot 1 = \alpha_2 = \lambda_2$$



PCA

Η ανάλυση κύριων συνιστωσών(ΑΚΣ). Είναι μία από τις βασικές μεθόδους μείωσης των διαστάσεων των δεδομένων, που βασίζεται στα στατιστικά χαρακτηριστικά τους.

Με την ΑΚΣ επιδιώκεται η περιγραφή των ανυσμάτων **x** που περιγράφουν τα πρότυπα με άλλα ανύσματα **y** ίσων ή λιγότερων διαστάσεων των οποίων οι γραμμές του **y** θα έχουν μηδενική ετεροσυσχέτιση.

PCA

Η μέθοδος Karhunen–Loève

Έστω για ένα σύνολο *K* ανυσμάτων σε χώρο *ν*-διαστάσεων διαφόρων χαρακτηριστικών και το $\boldsymbol{x}_{\kappa}$ είναι το διάνυσμα που δείχνεται από την σχέση:

$$\boldsymbol{x}_{\boldsymbol{\kappa}} = \left[x_{1\kappa}, x_{2\kappa}, \ldots, x_{\nu\kappa}, \ldots, x_{N\kappa}\right]^T$$

Τότε, ο πίνακας συνδιασποράς $\boldsymbol{C}_{\boldsymbol{x}}$ όλων των ανυσμάτων του συνόλου εκπαίδευσης δίνεται από τη σχέση:

$$C_x = E[(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^T] = \frac{1}{K}\sum_{\kappa=1}^{K}(\boldsymbol{x}_{\boldsymbol{\kappa}}-\boldsymbol{\mu}_{\boldsymbol{x}})(\boldsymbol{x}_{\boldsymbol{\kappa}}-\boldsymbol{\mu}_{\boldsymbol{x}})^T = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & . & . & \sigma_{1\nu} & . & . & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & . & . & \sigma_{2\nu} & . & . & \sigma_{2N} \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ \sigma_{\nu 1} & \sigma_{\nu 2} & . & . & \sigma_\nu^2 & . & . & \sigma_{\nu N} \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ \sigma_{N1} & \sigma_{N2} & . & . & \sigma_{N\nu} & . & . & \sigma_N^2 \end{bmatrix}$$

όπου $\boldsymbol{\mu}_{\boldsymbol{x}} = E[\boldsymbol{x}]$

PCA

Η μέθοδος Karhunen–Loève

Για να επιτύχουμε αυτό, αναζητούμε ένα τέτοιο μετασχηματισμό $\boldsymbol{W}$ που θα εκφράζεται από ένα πίνακα NxN, ώστε τα διανύσματα $y = W^T \cdot x$ να έχουν πίνακα συνδιασποράς $\boldsymbol{C_y}$ με την ακόλουθη μορφή: $det\left(C_x - \lambda I\right)$

$$C_y = \begin{bmatrix} \lambda_1 & 0 & . & . & 0 & . & . & 0 \\ 0 & \lambda_2 & . & . & 0 & . & . & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & 0 & . & . & \lambda_v & . & . & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & 0 & . & . & 0 & . & . & \lambda_N \end{bmatrix} = \Lambda \ , \ \lambda_1 > \lambda_2 > \ldots > \lambda_v > \ldots > \lambda_N$$

Ακόμη ισχύει ότι $\mathbf{C_y} = E[yy^T] = E[\mathbf{W}^T\mathbf{x}(\mathbf{W}^T\mathbf{x})^T] = E[\mathbf{W}T\mathbf{xx}^T\mathbf{W}] = \mathbf{W}^T E[\mathbf{xx}^T]\mathbf{W} = \mathbf{W}^T\mathbf{C}_x\mathbf{W}$

PCA

Η μέθοδος Karhunen–Loève

Είναι φανερό ότι η τιμή $\lambda_v = \sigma_v^{'2}$ είναι η διασπορά του *y* κατά τον *v* άξονα, ο οποίος εκφράζεται από την *v* στήλη του πίνακα *W*.

Η στήλη w_v είναι "αποτελεσματικό" χαρακτηριστικό ενώ η τιμή λ_v αξιολογεί τη σπουδαιότητά του. Ο πίνακας C_x είναι συμμετρικός και ο πίνακας Λ διαγώνιος.

Η σχέση $W^T C_x W = \Lambda$ ικανοποιείται αν $w_1,\ldots,w_N$ είναι τα ιδιοανύσματα του πίνακα C_x, $W=[w_1,\ldots,w_N]$. Οι ιδιοτιμές του C_x είναι οι τιμές $\lambda_1,\ldots,\lambda_N$.

Σε περίπτωση που απαιτείται μείωση του πλήθους N των ιδιοδιανυσμάτων μπορούμε να επιλέξουμε τα M από αυτά ($M<N$) με τις μεγαλύτερες αντίστοιχες ιδιοτιμές, αποδεχόμενοι ένα μέσο τετραγωνικό σφάλμα

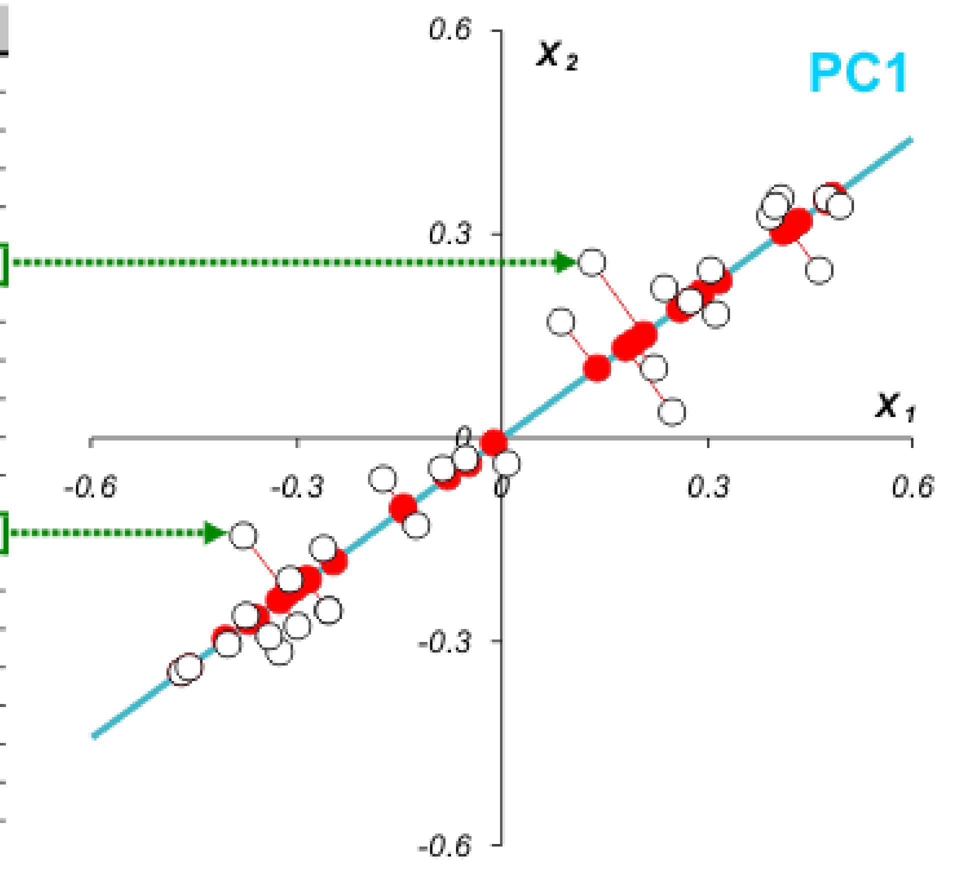
$$\overline{\varepsilon^2}(m) = \sum_{v=M+1}^{N} \lambda_v$$

PCA

Η μέθοδος Karhunen–Loève

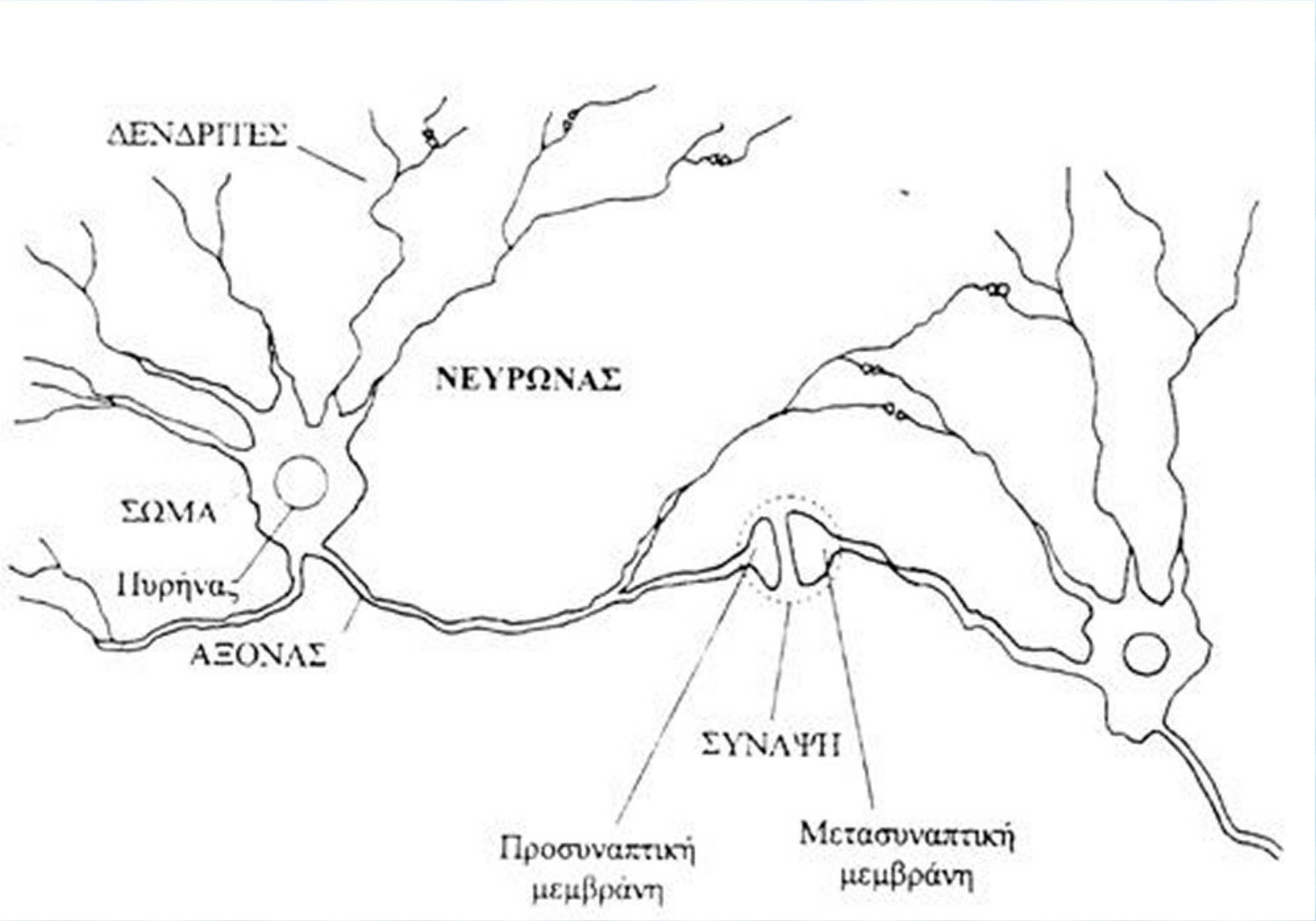
Στην παρακάτω εικόνα φαίνονται προβολές των εισόδων πάνω στην κύρια συνιστώσα.

	X_1	X_2
1	0.407	0.353
2	0.475	0.355
3	-0.088	-0.045
4	0.394	0.325
5	0.274	0.202
6	0.131	0.258
7	-0.053	-0.031
8	-0.124	-0.128
9	-0.469	-0.344
10	0.088	0.171
11	-0.261	-0.162
12	0.401	0.341
13	-0.376	-0.143
14	-0.251	-0.255
15	-0.325	-0.316
16	0.464	0.248
17	-0.310	-0.207
18	0.307	0.247
19	-0.399	-0.303
20	-0.253	-0.253
21	-0.341	-0.291



Νευρωνικά δίκτυα

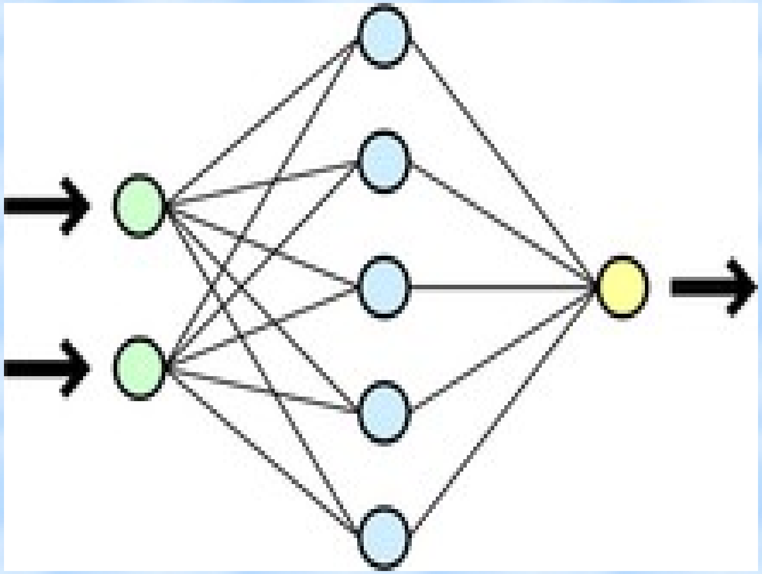
Βιολογικά νευρονικά δίκτυα



Νευρωνικά δίκτυα

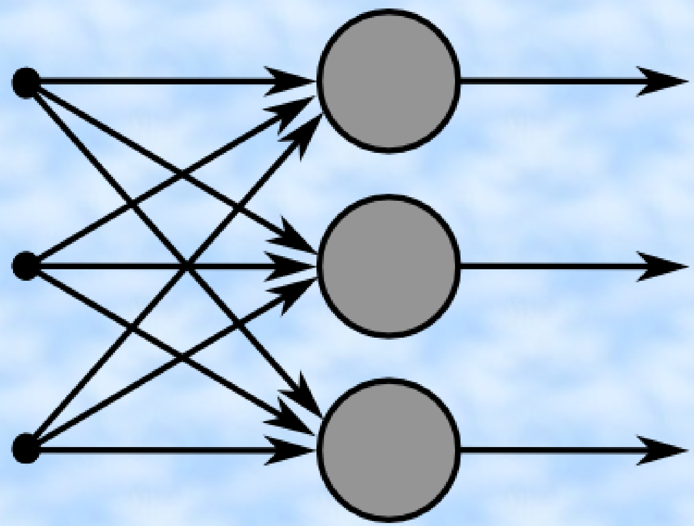
Τεχνητά μαθηματικά μοντέλα νευρωνικών δικτύων

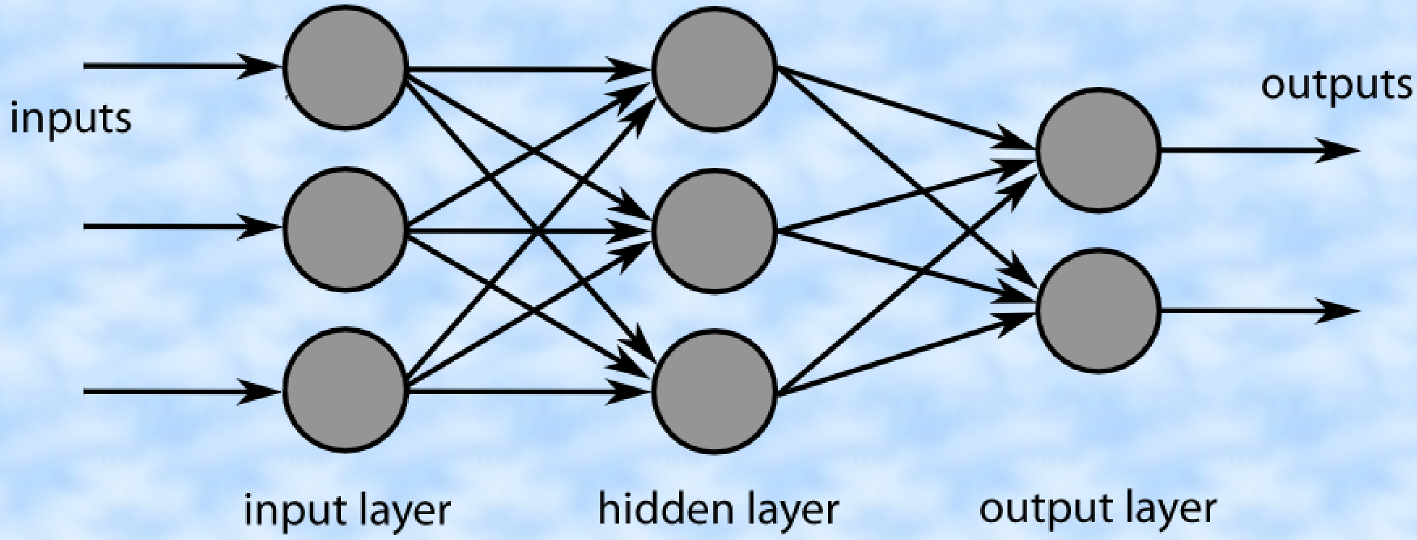
- Εκπαίδευση με επόπτη
- Εκπαίδευση χωρίς επόπτη
- Ενισχυτική μάθηση



Νευρωνικά δίκτυα

Ενός και πολλών επιπέδων





GHA

Αλγόριθμος του Χεμπ

Το πιο απλό μοντέλο του νευρώνα του Hebb περιγράφεται με μια γραμμική συνάρτηση με πολλαπλούς εισόδους x_1, x_2,..., x_m μέσω των συναπτικών βαρών w_1, w_2,..., w_m. Η έξοδος εκφράζεται με τον τυπο:

$$y=\sum_{i=1}^{m} w_i x_i$$

Τα βάρη μεταβάλλονται στο χρόνο συνεχώς, καθώς αυξάνονται δραματικά εάν η είσοδος x_i συμπίπτει με τη έξοδο y.

$$w_{ij}(n+1)=w_{ij}(n)+\gamma x_j(n)\, y_i(n)$$

Όπου γ είναι ένας συντελεστής που καθορίζει το ρυθμό μάθησης και το n είναι διακριτή χρονική στιγμή.

GHA

Γενικευμένος αλγόριθμος του Χεμπ

Έστω ένα μονού επιπέδου δίκτυο με εισόδους ένα N-διάστατο διάνυσμα με στήλη x και έναν πίνακα βαρών W με διάσταση MxN και εξόδους ένα διάνυσμα στήλη y=Wx με διάσταση M όπου M<N. Έστω επίσης οι τιμές του x που παράγονται από ένα στατικό τυχαίο διάνυσμα μιας στοχαστικής διεργασίας με πίνακα συσχέτισης $C_x=E[xx^T]$. Κατά συνέπεια θεωρείται πως τα x και y είναι συναρτήσεις του χρόνου και ο W θα είναι επίσης χρονικά εξαρτώμενος. Η σχέση που περιγράφει το GHA είναι:

$$W_{ji}(t+1)=w_{ji}(t)+\gamma(t)\left[y_i(t)\,x_i(t)-\sum_{k=1}^{j} w_{kj}(t)\,y_k(t)\right]$$

για i=1,2,... m και για j=1,2,...k

GHA

Γενικευμένος αλγόριθμος του Χεμπ

Ο GHA μπορεί να γραφτεί με τη μορφή:

$$\Delta W_{ji} = \gamma(n) y_j(n)[x_i(n) - \sum_{k=1}^{j-1} W_{ki}(n) y_k(n)] + \gamma(n) y_j^2(n) W_{ji}(n)$$

Η διαμορφωμένη μορφή της εισόδου δίνεται από τη σχέση:

$$x'_i(n) = x_i(n) - \sum_{k=1}^{j} w_{kj}(n) y_k(n)$$

Το κύριο στοιχείο στη διαμορφωμένη είσοδο είναι το i ιδιοδιάνυσμα της αρχικής εισόδου. Όταν εφαρμόζεται ο αλγόριθμος Oja στη διαμορφωμένη είσοδο έχει ως αποτέλεσμα η i έξοδος να μαθαίνει το i ιδιοδιάνυσμα.

DCT-2

Ο δισδιάστατος μετασχηματισμός συνημιτόνου (2D-DCT: Discrete Cosine Transform) αποτελεί την βάση για συμπίεση μη δυαδικών εικόνων με αποδεκτή απώλεια πληροφορίας. Η τυποποίηση JPEG βασίζεται στις πρώτες εκδόσεις της στον 2D-DCT.

Για ένα δισδιάστατο διακριτό σήμα $x(n_1,n_2)$ με μήκη N_1, N_2, οι προβολές του στα παραπάνω ορθοκανονικά ανύσματα βάσης είναι:

$$c_{k_1 k_2}=\frac{2}{\sqrt{N_1 N_2}}\sum_{n_2=0}^{N_2-1}\sum_{n_1=0}^{N_1-1} x[n_1,n_2]\cdot\cos\left(\frac{\pi k_1(2n_1+1)}{2}N_1\right)\cdot\cos\left(\frac{\pi k_2(2n_2+1)}{2N_2}\right)$$

Το σήμα $x[n_1,n_2]$ ανακτάται σύμφωνα με τη σχέση:

$$x(n_1,n_2)=\sum_{n_2=0}^{N_2-1}\sum_{n_1=0}^{N_1-1} c_{k_1 k_2}\cdot g_{k_1 k_2}(n_1,n_2)$$

PCA vs DCT-2

Εφαρμόζουμε τις δύο τεχνικές για υπολογισμό του μέσου τετραγωνικού σφάλματος mse (e^2) σε μια εικόνα Lenna, που χρησιμοποιείται για δοκιμές αλγορίθμων σε πολλά επιστημονικά έργα και έχει διαστάσεις 512x512.

Samples	64(8X8)	16(4X4)	4(2X2)
PCA	27.08	9.84	22.497
DCT-2	11.826	22.90	62.24

Για την περίπτωση PCA επιλέχθηκαν μισές κύριες συνιστώσες και για DCT-2 μισές συχνότητες.

PCA vs DCT-2

PCA

1 – αρχική εικόνα, 2 - συμπιεσμένη εικόνα, 3 – κύριες συνιστώσες(προβολή)

DCT-2

1 – αρχική εικόνα, 2 - συμπιεσμένη εικόνα, 3 – κύριες συνιστώσες(προβολή)