ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

Ψηφιακή Επεξεργασία Εικόνων

Ενότητα: 2-Δ ορθομοναδιαίοι μετασχηματισμοί

Καθηγητής Γεώργιος Τζιρίτας

Τμήμα Επιστήμης Υπολογιστών

Κεφάλαιο 8

2-Δ ορθομοναδιαίοι μετασχηματισμοί

Αντικείμενο αυτού του κεφαλαίου είναι οι διακριτοί μετασχηματισμοί, που χρησιμεύουν στην εξαγωγή ορισμένων χαρακτηριστικών από τις εικόνες, όπως επίσης στην παράσταση στο πεδίο των συχνοτήτων, και γι' αυτό ακριβώς το λόγο χρησιμοποιούνται στη συμπίεση των εικόνων. Οι 2-Δ μετασχηματισμοί γίνονται σε ορθογώνια ή τετράγωνα μπλοκ. Για λόγους απλότητας θα περιορισθούμε εδώ στην περίπτωση των τετράγωνων μπλοκ.

Ας θεωρήσουμε ότι ένα τετράγωνο μπλοκ παριστάνεται από ένα τετραγωνικό πίνακα X, διαστάσεων $N \times N$. Ας είναι E_N ο διανυσματικός χώρος των τετραγωνικών πινάκων διαστάσεων $N \times N$, με εσωτερικό γινόμενο ορισμένο κατά τον ακόλουθο τρόπο

$$\langle A, B\rangle = \text{ιχν}[B^{\dagger}A] = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} a(m,n)b^{\star}(m,n), \tag{8.1}$$

όπου $A \in E_N, B \in E_N$, και $B^{\dagger}$ είναι ο ανάστροφος και συζυγής πίνακας του B. Ένας γραμμικός μετασχηματισμός είναι μία γραμμική απεικόνιση του E_N στον E_N, που ορίζεται ως εξής

$$\begin{gathered} E_N \ni X \xrightarrow{\mathcal{T}} Y \in E_N \\ y(m,n) = \sum_{i=0}^{N-1}\sum_{k=0}^{N-1} t^{*}(m,n;i,k)x(i,k). \end{gathered} \tag{8.2}$$

Σύμφωνα με τον ορισμό του εσωτερικού γινομένου (8.1) μπορούμε να γράψουμε

$$y(m,n) = \langle X, \Phi(m,n)\rangle, \tag{8.3}$$

όπου το τυπικό στοιχείο του πίνακα $\Phi(m,n)$ είναι $\{t(m,n;i,k)\}$. Τα στοιχεία του πίνακα Y ονομάζονται συντελεστές του μετασχηματισμού. Οι πίνακες $\Phi(m,n)$ συνιστούν τη βάση του μετασχηματισμού.

Υποθέτουμε ότι ο μετασχηματισμός $\mathcal{T}$ είναι αντιστρέψιμος, δηλαδή ότι υπάρχει μετασχηματισμός, $\mathcal{T}^{-1}$, τέτοιος ώστε

$$\mathcal{T}\mathcal{T}^{-1} = \mathcal{T}^{-1}\mathcal{T} = \mathcal{I}, \tag{8.4}$$

όπου $\mathcal{I}$ είναι ο μοναδιαίος μετασχηματισμός του διανυσματικού χώρου E_N. Ο ανάστροφος και συζυγής μετασχηματισμός, $\mathcal{T}^{\dagger}$, έχει τυπικό στοιχείο $\{t^{*}(i,k;m,n)\}$. Στη συνέχεια θα περιορισθούμε στους ορθομοναδιαίους μετασχηματισμούς, οι οποίοι είναι αντιστρέψιμοι, και ο αντίστροφος είναι ίσος με τον ανάστροφο και συζυγή

$$\mathcal{T}^{-1} = \mathcal{T}^{\dagger}. \tag{8.5}$$

Σύμφωνα με τους παραπάνω ορισμούς προκύπτει ότι η βάση ενός ορθομοναδιαίου μετασχηματισμού είναι ορθοκανονική

$$\langle \Phi(m,n), \Phi(m',n') \rangle = \delta(m-m', n-n'). \tag{8.6}$$

Μπορεί επίσης να αποδειχθεί ότι η βάση είναι πλήρης. Επομένως ένας ορθομοναδαίος μετασχηματισμός συνιστά μία ανάλυση σε μία πλήρη και ορθοκανονική βάση.

Ιδιαίτερο ενδιαφέρον έχουν οι διαχωρίσιμοι μετασχηματισμοί, για τους οποίους ισχύει

$$t(m,n;i,k) = t_1(m,i)t_2(n,k), \forall i,k,m,n \tag{8.7}$$

Σε αυτή την περίπτωση δύο τετραγωνικοί πίνακες, T_1 και T_2, αρκούν για τον ορισμό του μετασχηματισμού, που δίδεται με τη βοήθεια δύο γινομένων πινάκων

$$Y = T_1^* X T_2^{\dagger}. \tag{8.8}$$

Όσον αφορά το μετασχηματισμό $\mathcal{T}$, είναι ίσος με το γινόμενο Kronecker των δύο συζυγών πινάκων

$$\mathcal{T} = T_1^* \otimes T_2^*. \tag{8.9}$$

Το γινόμενο Kronecker δύο τετραγωνικών πινάκων A και B ορίζεται ως εξής

$$A \otimes B = \begin{bmatrix} a(0,0)B & \cdots & a(0,N-1)B \\ \vdots & & \vdots \\ a(N-1,0)B & \cdots & a(N-1,N-1)B \end{bmatrix} \tag{8.10}$$

Ο αντίστροφος μετασχηματισμός είναι επίσης διαχωρίσιμος, και εφόσον είναι ορθομοναδιαίος δίδεται ως ακολούθως

$$\mathcal{T}^{-1} = T_1^t \otimes T_2^t, \tag{8.11}$$

οπότε θα έχουμε

$$X = T_1^t Y T_2. \tag{8.12}$$

Η πλέον συνηθισμένη περίπτωση είναι όταν $T_1 = T_2 = T$, οπότε είναι

$$Y = T^* X T^{\dagger} \;\; \text{και} \;\; X = T^t Y T. \tag{8.13}$$

Ο αριθμός των πολλαπλασιασμών και προσθέσεων που απαιτούνται στην περίπτωση ενός διαχωρίσιμου μετασχηματισμού είναι της τάξης $2N^3$, σε σχέση με N^4 που είναι στη γενική περίπτωση.

Κάθε ορθομοναδιαίος μετασχηματισμός έχει την ιδιότητα της διατήρησης της ενέργειας

$$\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} y^2(m,n) = \| Y \|^2 = \langle \mathcal{T}X, \mathcal{T}X \rangle = \langle \mathcal{T}^{\dagger}\mathcal{T}X, X \rangle = \langle X, X \rangle = \| X \|^2 = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} x^2(m,n). \tag{8.14}$$

Στη συνέχεια δίνουμε τους πλέον χρησιμοποιούμενους ορθομοναδιαίους μετασχηματισμούς.

8.1 Διακριτός μετασχηματισμός Fourier

Το τυπικό στοιχείο του διακριτού μετασχηματισμού Fourier (DFT, Discrete Fourier Transform) είναι το εξής

$$t(m,n;i,k) = \frac{1}{N}\exp(\frac{2\pi j}{N}(im+kn)) \tag{8.15}$$

Από τον ορισμό είναι προφανές ότι ο DFT είναι διαχωρίσιμος, και επίσης είναι εύκολο να αποδειχθεί ότι είναι ορθομοναδιαίος. Ο DFT είναι ευρύτατα χρησιμοποιούμενος μετασχηματισμός, για τον οποίο υπάρχουν γρήγοροι αλγόριθμοι υπολογισμού. Χάρη στη διαχωρισιμότητα, ο 2-Δ DFT μπορεί να υπολογισθεί με χρήση του 1-Δ FFT (Fast Fourier Transform), οπότε ο αριθμός των πολλαπλασιασμών ανέρχεται σε $N^2\log_2 N$. Εφαρμόζοντας την ίδια μέθοδο με τον 1-Δ FFT κατ′ ευθείαν στο 2-Δ DFT, μπορεί να μειωθεί ακόμη κατά 25% ο αριθμός των πολλαπλασιασμών. Γι′ αυτό θέτουμε: $W_N = e^{-\frac{j2\pi}{N}}$ και γράφουμε:

$$\begin{aligned}
S_{00}(m,n) &= \sum_{i=0}^{N/2-1}\sum_{k=0}^{N/2-1} x(2i,2k)W_N^{2im+2kn}\\
S_{01}(m,n) &= \sum_{i=0}^{N/2-1}\sum_{k=0}^{N/2-1} x(2i,2k+1)W_N^{2im+2kn}\\
S_{10}(m,n) &= \sum_{i=0}^{N/2-1}\sum_{k=0}^{N/2-1} x(2i+1,2k)W_N^{2im+2kn}\\
S_{11}(m,n) &= \sum_{i=0}^{N/2-1}\sum_{k=0}^{N/2-1} x(2i+1,2k+1)W_N^{2im+2kn}
\end{aligned} \tag{8.16}$$

Μπορούμε τότε να γράψουμε

$$\begin{aligned}
y(m,n) &= S_{00}(m,n) + W_N^n S_{01}(m,n) + W_N^m S_{10}(m,n) + W_N^{m+n} S_{11}(m,n)\\
y(m+\tfrac{N}{2},n) &= S_{00}(m,n) + W_N^n S_{01}(m,n) - W_N^m S_{10}(m,n) - W_N^{m+n} S_{11}(m,n)\\
y(m,n+\tfrac{N}{2}) &= S_{00}(m,n) - W_N^n S_{01}(m,n) + W_N^m S_{10}(m,n) - W_N^{m+n} S_{11}(m,n)\\
y(m+\tfrac{N}{2},n+\tfrac{N}{2}) &= S_{00}(m,n) - W_N^n S_{01}(m,n) - W_N^m S_{10}(m,n) + W_N^{m+n} S_{11}(m,n)
\end{aligned} \tag{8.17}$$

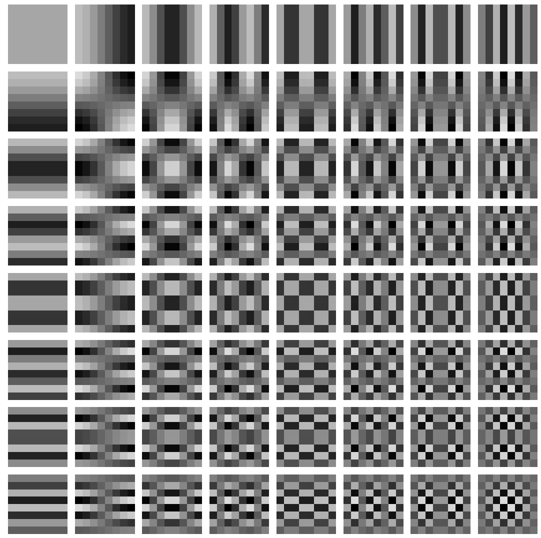
Το πλήθος υποδιαιρέσεων αυτού του τύπου είναι $\log_2 N$, και για κάθε τέσσερεις συντελεστές του μετασχηματισμού απαιτούνται τρείς πολλαπλασιασμοί. Επομένως ο αριθμός των πολλαπλασιασμών θα είναι $\frac{3N^2}{4}\log_2 N$.

8.2 Διακριτός μετασχηματισμός συνημιτόνου

Το τυπικό στοιχείο του διακριτού μετασχηματισμού συνημιτόνου (DCT, Discrete Cosine Transform) είναι το εξής

$$t(m,n;i,k) = \frac{c(m)c(n)}{N}\cos\frac{(2i+1)m\pi}{2N}\cos\frac{(2k+1)n\pi}{2N}, \tag{8.18}$$

όπου $c(m) = \begin{cases} 1 & m=0 \\ \sqrt{2} & m\neq 0 \end{cases}$. Ο DCT είναι ένας διαχωρίσιμος ορθομοναδιαίος μετασχηματισμός, που χρησιμοποιείται ευρύτατα για τη συμπίεση των εικόνων. Ο μετασχηματισμός συνημιτόνου είναι ιδιαίτερα αποτελεσματικός στη συμπίεση γιατί προσεγγίζει τις επιδόσεις του βέλτιστου γραμμικού μετασχηματισμού, εφόσον τα δεδομένα είναι ισχυρά συσχετισμένα.



Σχήμα 8.1: Η βάση του 2-Δ διακριτού μετασχηματισμού συνημιτόνου

Ο διακριτός μετασχηματισμός συνημιτόνου συνδέεται άμεσα με το διακριτό μετασχηματισμό Fourier. Ας είναι

$$x_1(m,n) = \begin{cases} x(m,n) & 0 \le m < N, 0 \le n < N \\ x(m, 2N-1-n) & 0 \le m < N, N \le n < 2N \\ x(2N-1-m, n) & N \le m < 2N, 0 \le n < N \\ x(2N-1-m, 2N-1-n) & N \le m < 2N, N \le n < 2N \end{cases} \tag{8.19}$$

και $y_1(m,n)$ ο διακριτός μετασχηματισμός Fourier του $x_1(m,n)$. Τότε αποδεικνύεται ότι, αν είναι $y(m,n)$ ο διακριτός μετασχηματισμός συνημιτόνου του $x(m,n)$, ισχύει

$$y(m,n) = \frac{1}{4} c(m) c(n) W_{2N}^{\frac{m+n}{2}} y_1(m,n), 0 \le m < N, 0 \le n < N \tag{8.20}$$

Οπότε για τον υπολογισμό του 2-Δ DCT μπορεί να χρησιμοποιηθεί ο 2-Δ FFT με διπλάσιο αριθμό σημείων ανά διάσταση. Ωστόσο υπάρχουν γρήγοροι αλγόριθμοι υπολογισμού που δεν απαιτούν τη χρήση του FFT με μικρότερο τελικά κόστος.

8.3 Διακριτός μετασχηματισμός ημιτόνου

Το τυπικό στοιχείο του διακριτού μετασχηματισμού ημιτόνου (DST, Discrete Sine Transform) είναι το εξής

$$t(m,n;i,k) = \frac{2}{N+1} \sin \frac{\pi(i+1)(m+1)}{N+1} \sin \frac{\pi(k+1)(n+1)}{N+1} \tag{8.21}$$

Είναι επίσης διαχωρίσιμος και ορθομοναδιαίος, και υπάρχουν γρήγοροι αλγόριθμοι για τον υπολογισμό του.

8.4 Διακριτός μετασχηματισμός Walsh-Hadamard

Το τυπικό στοιχείο του διακριτού μετασχηματισμού Walsh-Hadamard (WHT) ορίζεται με τις δυαδικές τιμές ± 1, αν αγνοήσουμε τον παράγοντα κανονικοποίησης $1/N$. Ο υπολογισμός επομένως του WHT δεν απαιτεί πολλαπλασιασμούς. Το τυπικό στοιχείο είναι το εξής

$$t(m,n;i,k) = \frac{1}{N}(-1)^{p(m,n;i,k)} \tag{8.22}$$

όπου $N = 2^l$ και $p(m,n;i,k) = \sum_{r=0}^{l-1}(i_r m_r^{(-)} + k_r n_r^{(-)})$. i_r (αντίστοιχα k_r) είναι οι δυαδικές παραστάσεις των i (αντίστοιχα k). Τα $m_r^{(-)}$ και $n_r^{(-)}$ περιλαμβάνουν επίσης μία αντιστροφή των bits, κατά τον ακόλουθο τρόπο

$$\begin{array}{lcl} m_0^{(-)} & = & m_{l-1} \\ m_1^{(-)} & = & m_{l-1} + m_{l-2} \\ \vdots & \vdots & \vdots \\ m_{l-1}^{(-)} & = & m_1 + m_0 \end{array} \tag{8.23}$$

Ο WHT είναι διαχωρίσιμος και ορθομοναδιαίος, και υπάρχουν γρήγοροι αλγόριθμοι για τον υπολογισμό του.

Άσκηση

Ας είναι $r(m,n)$ η συνάρτηση συμμεταβλητότητας ενός 2-Δ σήματος $x(m,n)$. Ο βέλτιστος γραμμικός ορθομοναδιαίος $N \times N$ μετασχηματισμός δίδεται από τη λύση των παρακάτω εξισώσεων:

$$\sum_{j=0}^{N-1} \sum_{l=0}^{N-1} r(i-j, k-l) t(m,n;j,l) = \lambda(m,n) t(m,n;i,k), \quad \text{για} \;\; 0 \leq i,k,m,n < N$$

Αν η συνάρτηση συμμεταβλητότητας είναι διαχωρίσιμη

$$r(m,n) = r_1(m) r_2(n),$$

τότε και ο μετασχηματισμός είναι διαχωρίσιμος

$$t(m,n;i,k) = t_1(m,i) t_2(n,k)$$

και προκύπτει ως εξής:

$$\sum_{j=0}^{N-1} r_1(i-j) t_1(m,j) = \lambda_1(m) t_1(m,i), \quad \text{για} \;\; 0 \leq m,i < N$$

$$\sum_{l=0}^{N-1} r_2(k-l) t_2(n,l) = \lambda_2(n) t_2(n,k), \quad \text{για} \;\; 0 \leq n,k < N$$

Δίδεται η ακόλουθη συνάρτηση συμμεταβλητότητας

$$r(m,n) = \sigma^2 \rho^{|m|+|n|}, \quad \text{με} \;\; 0 < \rho < 1$$

1. Να ευρεθεί ο 2×2 μετασχηματισμός που ικανοποιεί τις παραπάνω σχέσεις.
2. Να συγκριθεί το αποτέλεσμα με τους μετασχηματισμούς Fourier και συνημιτόνου.

Σημειώματα

Σημείωμα αναφοράς

Σημείωμα Αδειοδότησης

Διατήρηση Σημειωμάτων

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Κρήτης» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.