

Aspectos Teóricos das Transformadas de Imagens

WILLIAM ROBSON SCHWARTZ, HÉLIO PEDRINI

Universidade Federal do Paraná, Departamento de Informática
81531-990, Curitiba-PR, Brasil
{william, helio}@inf.ufpr.br

Resumo. Este trabalho aborda as principais transformadas de imagens baseadas em transformações de coordenadas, as quais executam uma rotação multidimensional sobre as coordenadas de um vetor. Os conceitos dessas transformações são apresentados, os quais normalmente são omitidos em referências de processamento de imagens. A transformada discreta de Fourier, amplamente utilizada em filtragem de imagens, descrição e classificação de objetos, tem seus conceitos e propriedades explorados. Resultados que possibilitam a criação da transformada rápida de Fourier são descritos. Algumas transformadas derivadas da transformada de Fourier, como a transformada discreta do cosseno, discreta do seno e a de Hartley são apresentadas.

1 Introdução

Este trabalho tem como objetivo apresentar um subconjunto das transformadas de imagens, que são baseadas nas transformações de coordenadas. A teoria dessas transformações não é facilmente encontrada em referências de processamento de imagens, que mostram, em sua maioria, apenas seus resultados e suas aplicações. O objetivo deste trabalho é apresentar sua teoria, encontrada na álgebra linear, e mostrar algumas de suas propriedades.

As transformações de coordenadas são um subconjunto das transformações lineares. Seus resultados podem ser interpretados como rotação multidimensional das coordenadas de um vetor, portanto, tem-se as mesmas informações, mas vistas sob outro ponto de vista, o que às vezes é importante para se ter uma melhor interpretação de uma imagem.

Essas transformadas foram utilizadas inicialmente na área de processamento digital de sinais, com seus casos unidimensionais (utilizando um vetor como entrada), e depois foram estendidas para o caso bidimensional para serem utilizadas em processamento digital de imagens, tendo como entrada uma imagem.

As imagens que são utilizadas estão normalmente no domínio espacial. Quando alguma transformação de coordenadas é aplicada, ela passa para o domínio de frequência, mostrando suas variações no domínio espacial, ou seja, se a imagem apresenta variações acentuadas no domínio espacial, ela contém componentes com alta frequência no domínio de frequência. Imagens no domínio de frequência podem facilitar a análise de resultados em algumas aplicações.

O trabalho é organizado como segue. A seção 2 apresenta conceitos da álgebra linear utilizados para definição das transformadas. A seção 3 descreve as transformadas de imagens, utilizando os conceitos apresentados na seção 2.

A seção 4 aborda a transformada discreta de Fourier e apresenta suas principais propriedades. A seção 5 mostra algumas transformadas que derivam da transformada de Fourier. Finalmente, na seção 6 são descritas as conclusões obtidas a partir deste trabalho.

2 Transformações de Coordenadas

Esta seção apresenta os conceitos da álgebra linear utilizados para o desenvolvimento da teoria das transformadas discretas de imagens, mostrando os resultados necessários para definição das transformações de coordenadas. As provas dos teoremas apresentados nesta seção são encontradas em [1].

Definição 2.1 (vetor) Um vetor \mathbf{v} com N componentes é representado como mostra a equação 1, v_i pertence ao corpo numérico \mathfrak{F} . O vetor mostrado nesta equação é chamado de vetor linha com N componentes, e \mathbf{v}^T é chamado vetor coluna

$$\mathbf{v} = [v_0 \ v_1 \ v_2 \ \dots \ v_{N-1}] \quad (1)$$

Definição 2.2 (matriz) A matriz \mathbf{A} com M linhas e N colunas, denotada matriz $M \times N$, é representada como mostra a equação 2, $a_{i,j} \in \mathfrak{F}$, $i = 0, 1, \dots, M-1$ e $j = 0, 1, \dots, N-1$

$$\mathbf{A} = \begin{bmatrix} a_{0,0} & a_{0,1} & \dots & a_{0,N-1} \\ a_{1,0} & a_{1,1} & \dots & a_{1,N-1} \\ \vdots & \vdots & & \vdots \\ a_{M-1,0} & a_{M-1,1} & \dots & a_{M-1,N-1} \end{bmatrix} \quad (2)$$

Definição 2.3 (conjugado complexo de uma matriz)

Seja uma matriz \mathbf{A} com componentes pertencentes ao conjunto dos complexos (\mathbb{C}). O conjugado complexo de \mathbf{A}

é obtido substituindo cada componente pelo seu conjugado complexo, a matriz obtida é denotada por \mathbf{A}^* .

Definição 2.4 (matriz unitária) Uma matriz quadrada \mathbf{U} é unitária se, e somente se, $\mathbf{U}^{*\mathbf{T}}\mathbf{U} = \mathbf{I}$, $\mathbf{U}^{*\mathbf{T}}$ é chamada de adjunta de \mathbf{U} , onde \mathbf{I} é a matriz identidade.

Definição 2.5 (matriz ortogonal) Uma matriz unitária \mathbf{U} com todos seus componentes pertencentes aos reais é chamada de matriz ortogonal.

Teorema 2.1 Uma matriz \mathbf{U} é unitária se, e somente se, suas colunas (linhas) são mutuamente ortogonais unitárias, ou seja, o produto interno entre as colunas (linhas) é 0 se as colunas (linhas) são diferentes e igual a 1 se as colunas (linhas) são iguais.

Definição 2.6 (espaço vetorial linear sobre \mathfrak{F}) Um conjunto não vazio $V_n(\mathfrak{F})$ de vetores com n componentes pertencentes ao corpo numérico \mathfrak{F} é denominado espaço vetorial linear sobre \mathfrak{F} , ou espaço vetorial sobre \mathfrak{F} , somente se satisfizer as seguintes condições:

- i. $c\mathbf{x} \in V_n(\mathfrak{F})$, onde $c \in \mathfrak{F}$ e $\mathbf{x} \in V_n(\mathfrak{F})$.
- ii. $\mathbf{x} + \mathbf{y} \in V_n(\mathfrak{F})$, onde $\mathbf{x}, \mathbf{y} \in V_n(\mathfrak{F})$.

Definição 2.7 (base de um espaço vetorial) O conjunto de vetores $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ pertencentes a um espaço vetorial $V_n(\mathfrak{F})$ é chamado de base de $V_n(\mathfrak{F})$ se todo vetor $\mathbf{w} \in V_n(\mathfrak{F})$ puder ser expresso unicamente como a combinação linear $\mathbf{w} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots + a_k\mathbf{v}_k$, e $a_1, a_2, \dots, a_k \in \mathfrak{F}$. Uma base composta por k vetores é representada por $B_{\{k\}}$.

Definição 2.8 (representação matricial de uma base) Seja $B_{\{k\}}$ uma base do espaço vetorial $V_n(\mathfrak{F})$, composta pelos vetores $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$. A equação 3 mostra a representação na forma matricial para base $B_{\{k\}}$, em que os vetores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ são colocados nas linhas da matriz \mathbf{B} , $k \times n$

$$\mathbf{B} = \begin{bmatrix} v_{1_0} & v_{1_1} & \dots & v_{1_{n-1}} \\ v_{2_0} & v_{2_1} & \dots & v_{2_{n-1}} \\ \vdots & \vdots & & \vdots \\ v_{k_0} & v_{k_1} & \dots & v_{k_{n-1}} \end{bmatrix} \quad (3)$$

Teorema 2.2 Sejam duas bases $B_{1\{n\}} = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ e $B_{2\{n\}} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ de um espaço vetorial $V_n(\mathfrak{F})$, relacionadas pela equação $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}\mathbf{S}$, e sejam os vetores \mathbf{u} e \mathbf{v} as coordenadas de um vetor arbitrário $\mathbf{x} \in V_n(\mathfrak{F})$ em relação às bases $B_{1\{n\}}$ e $B_{2\{n\}}$, respectivamente. Então \mathbf{u} e \mathbf{v} são relacionados pela transformação linear não singular $\mathbf{u}^{\mathbf{T}} = \mathbf{S}\mathbf{v}^{\mathbf{T}}$, denominada transformação de coordenadas. \mathbf{S} é uma matriz de ordem n , chamada de núcleo da transformação.

Como \mathbf{u} tem as coordenadas do vetor \mathbf{x} em relação à base $B_{1\{n\}}$ e \mathbf{v} tem as coordenadas do vetor \mathbf{x} em relação à base $B_{2\{n\}}$, \mathbf{x} pode ser representado como mostrado nas equações 4 e 5, enquanto \mathbf{u} e \mathbf{v} podem ser representados como mostrado nas equações 6 e 7. \mathbf{B}_1 e \mathbf{B}_2 são as representações matriciais das bases $B_{1\{n\}}$ e $B_{2\{n\}}$, respectivamente. As matrizes \mathbf{B}_1 e \mathbf{B}_2 admitem inversa porque são bases de $V_n(\mathfrak{F})$

$$\mathbf{x}^{\mathbf{T}} = \mathbf{B}_1\mathbf{u}^{\mathbf{T}} \quad (4)$$

$$\mathbf{x}^{\mathbf{T}} = \mathbf{B}_2\mathbf{v}^{\mathbf{T}} \quad (5)$$

$$\mathbf{u}^{\mathbf{T}} = \mathbf{B}_1^{-1}\mathbf{x}^{\mathbf{T}} \quad (6)$$

$$\mathbf{v}^{\mathbf{T}} = \mathbf{B}_2^{-1}\mathbf{x}^{\mathbf{T}} \quad (7)$$

Substituindo a equação 4 na equação 7, obtém-se a equação 8, que é uma transformação linear que expressa as coordenadas de \mathbf{x} com relação à base $B_{2\{n\}}$ diretamente em termos das coordenadas de \mathbf{x} com respeito à base $B_{1\{n\}}$. O núcleo da transformação, \mathbf{S} , utilizado pelo teorema 2.2, e seu inverso são mostrados pelas equações 9 e 10, respectivamente

$$\mathbf{v}^{\mathbf{T}} = (\mathbf{B}_2^{-1}\mathbf{B}_1)\mathbf{u}^{\mathbf{T}} \quad (8)$$

$$\mathbf{S} = \mathbf{B}_2^{-1}\mathbf{B}_1 \quad (9)$$

$$\mathbf{S}^{-1} = \mathbf{B}_1^{-1}\mathbf{B}_2 \quad (10)$$

Desde que as bases sejam conhecidas, pode-se criar uma transformação de coordenadas. Utiliza-se a equação 8 para transformar as coordenadas de um vetor que estão em relação à base $B_{1\{n\}}$ para coordenadas desse vetor em relação à base $B_{2\{n\}}$. A transformação inversa de 8 pode ser implementada, utilizando \mathbf{S}^{-1} como núcleo da transformada, como mostra a equação 11

$$\mathbf{u}^{\mathbf{T}} = \mathbf{S}^{-1}\mathbf{v}^{\mathbf{T}} \quad (11)$$

Definição 2.9 (transformação unitária) Se o núcleo da transformação de coordenadas for uma matriz unitária, esta transformação será denominada transformação unitária.

Definição 2.10 (transformação ortogonal) Se o núcleo da transformação de coordenadas for uma matriz ortogonal, esta transformação será denominada transformação ortogonal.

Teorema 2.3 Uma transformação de coordenadas $\mathbf{y}^{\mathbf{T}} = \mathbf{A}\mathbf{x}^{\mathbf{T}}$ mantém o tamanho de todos os vetores e a ortogonalidade entre eles se, e somente se, ela for uma transformação unitária ou ortogonal.

3 Transformadas de Imagens

Em processamento de imagens, o termo *transformada* é utilizado para designar classes distintas de transformações. Este trabalho trata de uma subclasse das transformadas, composta pelas transformações de coordenadas, as quais são um subconjunto das transformações lineares. As transformações de coordenadas consistem na rotação multidimensional das coordenadas de um ponto [2].

As transformadas são unidimensionais (1D) quando a entrada é um vetor, e bidimensionais (2D) quando a entrada é uma matriz. Em processamento de imagens é comum utilizar transformadas bidimensionais. É importante ressaltar que o termo unidimensional ou bidimensional está relacionado com a dimensionalidade do elemento de entrada, vetor ou matriz, no entanto, a transformação linear é executada de $\mathbb{C}^N \rightarrow \mathbb{C}^N$, sendo N o número de componentes do vetor ou da matriz.

As transformadas possuem representação vetorial e com somatórios, como mostrado na tabela 1. \mathbf{X} e \mathbf{x} contêm a entrada da transformada, \mathbf{Y} e \mathbf{y} contêm o seu resultado. Seus componentes são chamados de coeficientes da transformada. \mathbf{X} e \mathbf{Y} são matrizes $M \times N$, \mathbf{x} e \mathbf{y} são vetores com N componentes. As matrizes \mathbf{A} e \mathbf{B} são os núcleos da transformação. \mathbf{A} é uma matriz $N \times N$ e \mathbf{B} é representada por um conjunto composto por UV matrizes $M \times N$.

Unidimensional	Bidimensional
$\mathbf{y}^T = \mathbf{A}\mathbf{x}^T$ $y_u = \sum_{n=0}^{N-1} a_{u,n}x_n$	$\mathbf{Y} = \mathbf{B}\mathbf{X}$ $y_{u,v} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} b_{u,v;m,n}x_{m,n}$

Tabela 1: Representação vetorial e com somatórios para transformadas unidimensionais e bidimensionais.

A seção 3.1 trata do núcleo da transformada, denominado a partir de agora de núcleo da transformada. A seção 3.2 mostra a representação dos pixels de uma imagem como coordenadas de uma base de um espaço vetorial, o que possibilita a definição das transformadas de imagens como transformações de coordenadas. A seção 3.3 mostra as representações das transformadas de imagens.

3.1 Núcleo da Transformada

O núcleo de uma transformada é a matriz \mathbf{S} ou \mathbf{S}^{-1} da transformação de coordenadas mostrada pelas equações 9 e 10. Transformadas são transformações que alteram as coordenadas de um vetor em relação à base $B_{1\{n\}}$ para coordenadas desse vetor em relação à base $B_{2\{n\}}$ utilizando a equação 8 ou 11. A seguir é mostrado um exemplo de determinação do núcleo da transformada.

Exemplo 3.1 A transformada discreta de Fourier (descrita na seção 4) é uma transformação unitária de um vetor \mathbf{v} , com N componentes, da base canônica para base $B_{2\{N\}}$, que tem sua matriz mostrada a seguir

$$\mathbf{B}_2 = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega_1^1 & \omega_1^2 & \dots & \omega_1^{N-1} \\ 1 & \omega_2^1 & \omega_2^2 & \dots & \omega_2^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_{N-1}^1 & \omega_{N-1}^2 & \dots & \omega_{N-1}^{N-1} \end{bmatrix} \quad (12)$$

onde $\omega_u^n = e^{2\pi i n u / N}$

A obtenção do núcleo da transformada de Fourier é dada pela utilização da equação 9, portanto, o núcleo da transformada unidimensional de Fourier é

$$\mathbf{A} = \mathbf{B}_2^{-1}\mathbf{I} = \mathbf{B}_2^{-1} \quad (13)$$

Na seção 3.2 será mostrado que os elementos de uma imagem (os pixels) podem ser considerados como coordenadas de um ponto em relação à base canônica, portanto, não há a necessidade de utilizar diretamente a equação 9 para determinação do núcleo de uma transformada, o núcleo sempre será a matriz da base de destino ou sua inversa.

O núcleo de uma transformada unidimensional é representado por uma matriz \mathbf{A} com ordem N , que contém em suas linhas os vetores $\mathbf{a}_u = [a_{u_0} \ a_{u_1} \ \dots \ a_{u_{N-1}}]$. Essa matriz normalmente é representada graficamente por N planos cartesianos, cada plano representa um vetor da base e cada ponto é um componente do u -ésimo vetor. A figura 1 mostra a representação gráfica do núcleo da transformada discreta de cosseno (descrita na seção 5.2), para $N = 8$.

No caso bidimensional, o núcleo é formado por um conjunto composto de UV matrizes $M \times N$, onde tais matrizes podem ser representadas por uma imagem em tons de cinza composta por $U \times V$ blocos, sendo que cada bloco representa uma matriz $M \times N$. A figura 2 mostra o núcleo da DCT para $N = M = U = V = 8$.

3.2 Definição de Imagem

Inicialmente, uma imagem será considerada um vetor \mathbf{f} com T componentes. Agora será mostrado que o elementos do vetor \mathbf{f} , os pixels da imagem, podem ser considerados como coordenadas em relação à uma base $B_{\{T\}}$ de um espaço vetorial. A base $B_{\{T\}}$ tem sua matriz definida como mostra a equação 14

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \quad (14)$$

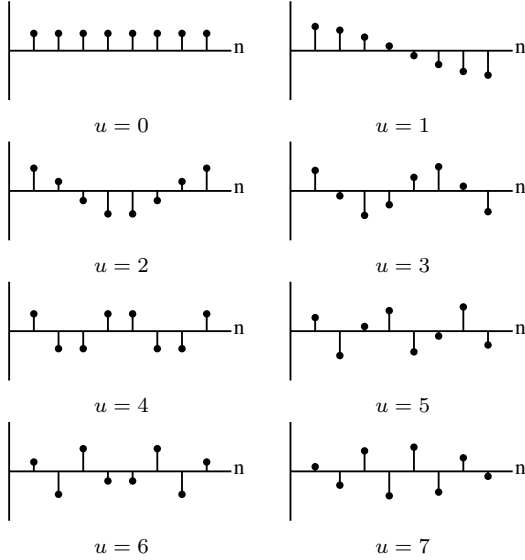


Figura 1: Representação gráfica do núcleo da transformada discreta do cosseno (DCT) para um vetor de entrada com 8 componentes.

A equação 15 mostra a combinação linear do vetor \mathbf{f} com os vetores da base $B_{\{T\}}$, resultando no vetor \mathbf{i} , que representa as coordenadas do vetor \mathbf{f} em relação à base $B_{\{T\}}$

$$i_m = \sum_{n=0}^{T-1} f_n B_{m,n}, \quad m = 0, 1, \dots, T-1 \quad (15)$$

Portanto, uma imagem pode ser considerada um ponto em um espaço vetorial gerado pela base $B_{\{T\}}$, denominada base canônica. Cada componente do vetor \mathbf{i} é uma coordenada desse ponto em relação a um eixo da base.

A transformação de coordenadas (unitária ou ortogonal) faz a mudança das coordenadas de uma imagem que estão em relação à base $B_{\{T\}}$ para coordenadas dessa imagem em relação à uma outra base $B_{2\{T\}}$, mantendo as informações e o número de dimensões.

Desde que as imagens são representadas por matrizes, a transformação de um vetor com $T = M \times N$ componentes para uma matriz \mathbf{J} , $M \times N$, pode ser feita pela decomposição de \mathbf{i} em M vetores com N componentes. O resultado é mostrado na equação 16

$$\mathbf{J} = \begin{bmatrix} i_0 & i_1 & \dots & i_{N-1} \\ i_N & i_{N+1} & \dots & i_{2N-1} \\ \vdots & \vdots & \ddots & \vdots \\ i_{(M-1)N} & i_{(M-1)N+1} & \dots & i_{(M \times N)-1} \end{bmatrix} \quad (16)$$

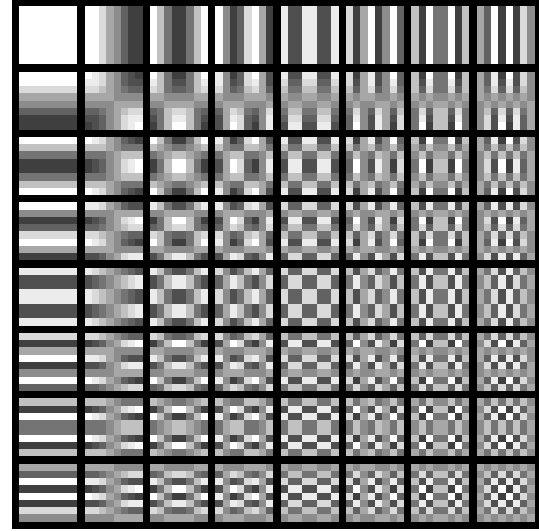


Figura 2: Representação gráfica do núcleo da transformada discreta do cosseno (DCT) para uma matriz de entrada com ordem 8.

3.3 Representação das Transformadas

Uma transformação de coordenadas para imagens pode ser escrita na forma mostrada pela equação 17, onde $j_{m,n}$ é um componente da matriz \mathbf{J} , $M \times N$, que contém a imagem de entrada. $a_{u,v;m,n}$ é o componente do núcleo da transformação, representado por \mathbf{A} , composto por um conjunto com UV matrizes $M \times N$. $o_{u,v}$ é um componente da matriz \mathbf{O} , $M \times N$, que contém a imagem resultante. A transformada inversa de 17 é mostrada na equação 18

$$o_{u,v} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} a_{u,v;m,n} j_{m,n} \quad (17)$$

$$\mathbf{O} = \mathbf{A}\mathbf{J}$$

$$j_{m,n} = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} a_{u,v;m,n}^{-1} o_{u,v} \quad (18)$$

$$\mathbf{J} = \mathbf{A}^{-1}\mathbf{O}$$

4 Transformada Discreta de Fourier (DFT)

Em processamento de imagens ou processamento de sinais, a série discreta de Fourier é chamada de transformada discreta de Fourier (DFT), o que pode causar alguma confusão, pois existe a transformada de Fourier que é diferente da série de Fourier [3], sendo que a segunda é periódica e a primeira não.

A DFT é uma transformação de coordenadas, que resulta em componentes pertencentes aos números complexos. Cada coeficiente é obtido pela combinação linear dos

elementos da entrada com o núcleo da transformada. O resultado da transformada de Fourier representa exatamente um período [3].

Os componentes complexos resultantes da DFT normalmente não são utilizados na forma $z = a + bi$. É comum utilizar a fase, representada pelo argumento de z (equação 19) ou o espectro de Fourier, que é o módulo de z (equação 20). A fase contém as informações essenciais sobre a estrutura da imagem, e o espectro de Fourier mostra o que a estrutura contém, mas sem nenhuma referência espacial [4]

$$\arg z = \tan^{-1} \left(\frac{b}{a} \right) \quad (19)$$

$$\text{mod } z = \sqrt{a^2 + b^2} \quad (20)$$

4.1 Transformada Unidimensional de Fourier

A transformada unidimensional de Fourier e sua inversa são mostradas pela equações 21 e 22, respectivamente, onde \mathbf{x} é o vetor de entrada e \mathbf{y} é o vetor de saída, ambos com N componentes, onde $u = 0, 1, \dots, N - 1$

$$y_u = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n e^{-2\pi i n u / N} \quad (21)$$

$$x_n = \sum_{u=0}^{N-1} y_u e^{2\pi i n u / N} \quad (22)$$

Substituindo u na equação 23, obtém-se os vetores da base para a transformada. Nota-se que o expoente de e é positivo e a equação 21 utiliza expoente negativo, isso porque é necessário utilizar a equação 8 para obtenção da equação de transformação de coordenadas da base canônica para a nova base, denominada base de Fourier

$$\mathbf{b}_u = \frac{1}{\sqrt{N}} [1 \quad \omega_u^1 \quad \omega_u^2 \quad \dots \quad \omega_u^{N-1}] \quad (23)$$

onde $\omega_u^n = e^{2\pi i n u / N}$

A matriz da base de Fourier é mostrada na equação 24. $B_{\{N\}}$ é realmente uma base, pois o conjunto formado pelos vetores $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{N-1}$ é ortogonal, garantido pela propriedade de ortogonalidade de senóides complexas [3]

$$\mathbf{B} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & \omega_1^1 & \omega_1^2 & \dots & \omega_1^{N-1} \\ 1 & \omega_2^1 & \omega_2^2 & \dots & \omega_2^{N-1} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & \omega_{N-1}^1 & \omega_{N-1}^2 & \dots & \omega_{N-1}^{N-1} \end{bmatrix} \quad (24)$$

4.2 Transformada Bidimensional de Fourier

A transformada bidimensional de Fourier é uma expansão da transformada unidimensional. As equações 25

e 26 mostram a transformada de Fourier e sua inversa respectivamente, para uma matriz de entrada \mathbf{X} e matriz resultante \mathbf{Y} , ambas com ordem N . A matriz \mathbf{Y} é denominada plano de freqüências, pois mapeia as variações da matriz de entrada

$$y_{u,v} = \frac{1}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x_{m,n} e^{-2\pi i (mu+nv)/N} \quad (25)$$

$$x_{m,n} = \sum_{v=0}^{N-1} \sum_{u=0}^{N-1} y_{u,v} e^{2\pi i (mu+nv)/N} \quad (26)$$

4.3 Propriedades da Transformada de Fourier

Nesta seção são mostradas algumas propriedades da transformada de Fourier, sendo estas bastante utilizadas em processamento digital de imagens [5, 6]. Nesta seção é considerada a transformada para uma matriz de entrada \mathbf{X} e matriz de saída \mathbf{Y} , ambas com ordem N e compostas pelos elementos $x_{m,n}$ e $y_{u,v}$, respectivamente.

Periodicidade A propriedade de periodicidade mostra que a transformada de Fourier tem período N , isto significa que $y_{u,v} = y_{(u+N),(v+N)}$. Portanto, são necessários somente N^2 coeficientes para determinar a transformada de Fourier para uma matriz de entrada com ordem N .

Simetria Conjugada A propriedade de simetria conjugada mostra que a magnitude da transformada está centrada na origem. Considerando o caso unidimensional, para o espectro de Fourier plotado no plano cartesiano, as magnitudes são refletidas em relação ao eixo y .

Se $x_{m,n}$ são componentes reais, então a transformada de Fourier apresenta simetria conjugada, ou seja, $y_{u,v} = y_{(-u,-v)}^*$.

Translação A propriedade da translação mostra que, ao multiplicar uma matriz no domínio espacial por um exponencial, a origem do plano de freqüências é transladada. O mesmo ocorre com a origem do plano espacial quando multiplica-se todos os coeficientes de Fourier por um exponencial. Essas propriedades são mostradas a seguir, \mathcal{F} é a transformada de Fourier aplicada sobre alguma entrada

$$\mathcal{F}(x_{m,n} e^{i2\pi(u_0 m + v_0 n)/N}) = y_{(u-u_0),(v-v_0)} \quad (27)$$

$$\mathcal{F}^{-1}(y_{u,v} e^{-i2\pi(m_0 u + n_0 v)/N}) = x_{(m-m_0),(n-n_0)} \quad (28)$$

A equação 27, mostra que ao multiplicar a matriz de entrada pelo exponencial $e^{i2\pi(u_0 m + v_0 n)/N}$, a origem do plano de freqüências é transladada para a coordenada (u_0, v_0) . A equação 28 mostra o mesmo resultado para a transformada inversa que, ao multiplicar a matriz resultante

pelo exponencial $e^{-i2\pi(m_0u+n_0v)/N}$, a origem do plano espacial é transladada para a coordenada (m_0, n_0) .

A propriedade de simetria conjugada mostra que a transformada de Fourier é simétrica em relação à origem, portanto, a origem do plano de frequências coincide com a origem da matriz de saída. Para o plano de frequências começar no início de um período, sua origem deve ser transladada para $u_0 = v_0 = N/2$, para isto, substitui-se $N/2$ na equação 27, obtendo-se $\mathcal{F}(x_{m,n}(-1)^{m+n})$.

Portanto, a DFT deve ser aplicada após a multiplicação de cada elemento da matriz \mathbf{X} por $(-1)^{m+n}$, facilitando a interpretação dos resultados. A figura 3 mostra o resultado da DFT utilizando ou não a translação da origem.

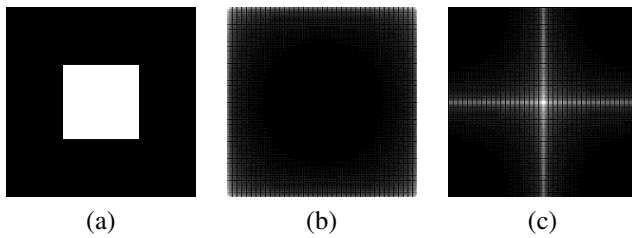


Figura 3: Transformada de Fourier utilizando ou não a translação da origem do plano de frequências. (a) Imagem original; (b) transformada sem utilizar a translação da origem; (c) DFT utilizando a translação da origem.

Um resultado importante da propriedade de translação é que o espectro de Fourier é invariante em relação à translação, pois $|y_{u,v}e^{-i2\pi(um_0+vn_0)/N}| = |y_{u,v}|$.

4.4 Conceito de Frequência em Imagens

Em sinais, como a corrente elétrica ou ondas sonoras, o conceito de frequência está relacionado com as variações da onda em um determinado intervalo de tempo, quanto maior for a variação do sinal em um intervalo pequeno de tempo, maior será sua frequência.

Em processamento de imagens, o conceito de frequência está relacionado com a variação da intensidade da função em relação ao espaço. Se a variação da intensidade dos pixels de uma região for grande, a frequência da região será alta.

A informação sobre quais frequências estão presentes em uma imagem são úteis, por exemplo, para segmentação das bordas de objetos, pois regiões de bordas são compostas por variações acentuadas na intensidade dos pixels, assim, a representação dessa imagem no domínio de frequência apresenta componentes com frequências altas.

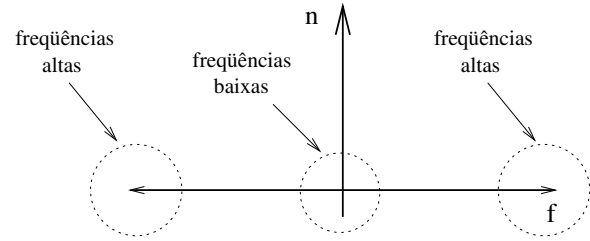


Figura 4: Plano de frequências para representação da transformada de Fourier unidimensional.

4.5 Interpretação da Transformada de Fourier

A transformada de Fourier mostra uma representação da função de entrada no domínio de frequência, ou seja, o resultado dessa transformada pode ser interpretado como o mapeamento das frequências presentes na entrada.

Para funções unidimensionais, as frequências podem ser representadas por um plano, como mostrado pela figura 4. Conforme ocorre o afastamento da origem acontece o aumento da frequência.

Para funções de entrada bidimensionais, a representação da transformada de Fourier pode ser feita através de uma superfície. Quando a função de entrada contém componentes com alta frequência, esta superfície apresentará elementos distantes da origem, caso contrário, a concentração dos elementos será em áreas próximas à origem.

4.6 Transformada Rápida de Fourier (FFT)

A transformada unidimensional de Fourier que foi abordada na seção 4.1, necessita de $O(N)$ operações com matemática complexa para determinação de cada coeficiente de Fourier, ou seja, para determinar todos os coeficientes, são necessárias $O(N^2)$ operações, se for implementada utilizando diretamente a equação 21.

Para diminuir o esforço computacional necessário para o cálculo da DFT, utiliza-se um método chamado de *transformada rápida de Fourier* (FFT), baseado na decomposição do núcleo da transformada em matrizes esparsas. O lema 4.1 mostra como o núcleo da DFT pode ser decomposto, sua prova pode ser encontrada em [7].

Lema 4.1 *Seja M um inteiro positivo. Então, o núcleo da transformada de Fourier, denotada matriz \mathbf{F} , com ordem $2M$, pode ser fatorada como mostra a equação 29*

$$\mathbf{F} = \frac{1}{\sqrt{2}} \mathbf{E}(\mathbf{N} \oplus \mathbf{N}) \mathbf{P} \quad (29)$$

onde $p_{m,n}^e \stackrel{\text{def}}{=} \delta(2m - n)$, $p_{m,n}^o \stackrel{\text{def}}{=} \delta(2m + 1 - n)$, para $m = 0, 1, \dots, M - 1$ e $n = 0, 1, \dots, 2M - 1$, $\mathbf{P} \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{P}^e \\ \mathbf{P}^o \end{pmatrix}$,

com ordem $2M$, onde \mathbf{P}^e e \mathbf{P}^o são matrizes $M \times 2M$. \mathbf{N} é o núcleo da DFT para uma transformada com M componentes, portanto, possui ordem M . $\mathbf{E} \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{I} & \mathbf{\Omega} \\ \mathbf{I} & -\mathbf{\Omega} \end{pmatrix}$, \mathbf{I} possui ordem M . A matriz $\mathbf{\Omega}$ possui ordem M , e é definida pela equação 30

$$\mathbf{\Omega} \stackrel{\text{def}}{=} \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & e^{-\pi i/M} & 0 & \dots & 0 \\ 0 & 0 & e^{-2\pi i/M} & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & e^{-(M-1)\pi i/M} \end{pmatrix} \quad (30)$$

O lema 4.1 mostra que o núcleo da DFT pode ser decomposto em matrizes esparsas, pois a matriz \mathbf{E} possui somente dois elementos diferentes de zero em cada linha ou coluna, e \mathbf{P} é uma matriz de permutação [7]. Para obtenção da transformada rápida de Fourier, as somas diretas são fatoradas recursivamente, como mostra o teorema 4.1.

Teorema 4.1 Se $N = 2^q$, $q \in \mathbb{N}$, o núcleo da DFT, \mathbf{F} , com ordem N , pode ser decomposto, como mostra a equação 31

$$\mathbf{F} = \frac{1}{\sqrt{N}} \mathbf{L}_0 \mathbf{L}_1 \dots \mathbf{L}_{q-1} \mathbf{P} \quad (31)$$

onde \mathbf{P} é a matriz definida pelo lema 4.1, com ordem N , e

$$\mathbf{L}_k = \overbrace{\mathbf{E} \oplus \dots \oplus \mathbf{E}}^{2^k \text{ vezes}} \quad (32)$$

cada matriz \mathbf{E} possui ordem $N/2^k$.

Portanto, a FFT é obtida decompondo o núcleo em várias matrizes esparsas, e os coeficientes de Fourier são calculados incrementalmente. Inicialmente, calcula-se a transformada para poucos pontos, e seus resultados são utilizados para a obtenção da transformada para pontos restantes, o cálculo da FFT necessita de $O(N \log_2 N)$ operações [8].

5 Transformadas Derivadas da DFT

Nesta seção são mostradas transformadas que derivam da transformada discreta de Fourier que são utilizadas em processamento de imagens [9, 10, 11].

5.1 Transformada Discreta de Hartley (DHT)

Para o cálculo da DFT, é necessária a criação de estruturas de dados para armazenamento da parte real e imaginária de um número, e a utilização de ferramentas que permitam cálculos com números complexos. A satisfação dessas condições gera a necessidade de maior espaço para armazenamento e aumento na capacidade de processamento. Como uma alternativa, utiliza-se a transformada de Hartley.

A transformada Hartley (DHT) é definida para o conjunto dos números reais. As equações 34 e 35 mostram a

transformada Hartley e sua inversa, para o caso bidimensional, com a matriz de entrada \mathbf{X} e matriz resultante \mathbf{Y} , ambas com ordem N .

$$\text{cas}(\theta) = \cos(\theta) + \text{sen}(\theta) \quad (33)$$

$$y_{u,v} = \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x_{m,n} \text{cas}\left(\frac{2\pi mu + 2\pi nv}{N}\right) \quad (34)$$

$$x_{m,n} = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} y_{u,v} \text{cas}\left(\frac{2\pi mu + 2\pi nv}{N}\right) \quad (35)$$

O núcleo da DHT é formado pela parte real e imaginária do núcleo da DFT, e a maioria das propriedades resultantes da transformada de Fourier também é obtida a partir desta transformada. A DHT apresenta ainda duas vantagens em relação à DFT: (i) o núcleo da transformada é igual ao de sua inversa, e (ii) utiliza somente computação real [12].

5.2 Transformada Discreta do Cosseno (DCT)

A núcleo da DCT é obtida a partir da parte real da base da transformada de Fourier, mas para continuar sendo uma base sofre algumas alterações [4]. A definição dos vetores componentes do núcleo da transformada unidimensional do cosseno, para um vetor de entrada com N componentes, é mostrada pela equação 36

$$b_{u,n} = \begin{cases} \frac{1}{\sqrt{N}}, & \text{se } u = 0 \\ \sqrt{\frac{2}{N}} \cos \frac{\pi(2n+1)u}{2N}, & \text{caso contrário} \end{cases} \quad (36)$$

As transformadas unidimensional do cosseno e sua inversa são mostradas pelas equações 38 e 39, respectivamente, para um vetor de entrada \mathbf{x} , composto por N elementos

$$k(\alpha) = \begin{cases} \frac{1}{\sqrt{N}}, & \text{se } \alpha = 0 \\ \sqrt{\frac{2}{N}}, & \text{caso contrário} \end{cases} \quad (37)$$

$$y_u = k(u) \sum_{n=0}^{N-1} x_n \cos\left(\frac{\pi(2n+1)u}{2N}\right) \quad (38)$$

$$x_n = \sum_{u=0}^{N-1} k(u) y_u \cos\left(\frac{\pi(2n+1)u}{2N}\right) \quad (39)$$

No caso bidimensional, as transformadas do cosseno e sua inversa são definidas pelas equações 42 e 43, respecti-

vamente, para uma matriz de entrada com ordem N

$$k(\alpha) = \begin{cases} \frac{1}{\sqrt{N}}, & \text{se } \alpha = 0 \\ \sqrt{\frac{2}{N}}, & \text{caso contrário} \end{cases} \quad (40)$$

$$\gamma(m, u) = \cos\left(\frac{\pi(2m+1)u}{2N}\right) \quad (41)$$

$$y_{u,v} = k(u)k(v) \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x_{m,n} \gamma(m, u) \gamma(n, v) \quad (42)$$

$$x_{m,n} = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} k(u)k(v) y_{u,v} \gamma(m, u) \gamma(n, v) \quad (43)$$

A transformada discreta do cosseno é real e ortogonal. A sua principal propriedade é a de compactar a energia em regiões próximas da origem para dados de entrada que apresentam alta correlação [4].

5.3 Transformada Discreta do Seno (DST)

A transformada do seno também é derivada da transformada de Fourier, mas seu núcleo não é diretamente a parte imaginária da DFT, ele é alterado para que continue sendo uma base. O núcleo da DST unidimensional, para um vetor de entrada com N componentes, pode ser obtido pela expansão da equação 44

$$b_{u,n} = \sqrt{\frac{2}{N+1}} \operatorname{sen}\left(\frac{\pi(u+1)(n+1)}{N+1}\right) \quad (44)$$

As transformadas unidimensional do seno e sua inversa são mostradas nas equações 45 e 46, respectivamente, para o vetor de entrada \mathbf{x} , composto por N elementos

$$y_u = \sqrt{\frac{2}{N+1}} \sum_{n=0}^{N-1} x_n \operatorname{sen}\left(\frac{\pi(n+1)(u+1)}{N+1}\right) \quad (45)$$

$$x_n = \sqrt{\frac{2}{N+1}} \sum_{u=0}^{N-1} y_u \operatorname{sen}\left(\frac{\pi(n+1)(u+1)}{N+1}\right) \quad (46)$$

No caso bidimensional, as transformadas do seno e sua inversa são definidas pelas equações 48 e 49, respectivamente, para uma matriz de entrada com ordem N

$$\phi(m, u) = \operatorname{sen}\left(\frac{\pi(m+1)(u+1)}{N+1}\right) \quad (47)$$

$$y_{u,v} = \frac{2}{N+1} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x_{m,n} \phi(m, u) \phi(n, v) \quad (48)$$

$$x_{m,n} = \frac{2}{N+1} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} y_{u,v} \phi(m, u) \phi(n, v) \quad (49)$$

6 Conclusões

Neste trabalho foram abordados conceitos da álgebra linear utilizados para a definição das transformadas de imagens. Desde que a teoria dessas transformadas é pouco abordada em referências de processamento de imagens, procurou-se apresentar seus principais aspectos.

Referências

- [1] Franz E. Hohn, *Elementary Matrix Algebra*, The Macmillan Company, 1958.
- [2] William K. Pratt, *Digital Image Processing*, John Wiley & Sons, Inc., Nova Iorque, Estados Unidos, 1978.
- [3] Simon Haykin and Barry van Veen, *Sinais e Sistemas*, Bookman, Porto Alegre, RS, Brasil, 2001.
- [4] Bernd Jähne, *Digital Image Processing*, Springer Verlag, Berlim, Alemanha, 5ª edição, 2002.
- [5] Guangyi Chen and Tien D. Bui, “Invariant Fourier-wavelet Description for Pattern Recognition,” *Pattern Recognition*, vol. 32, pp. 1083–1088, 1999.
- [6] Hubert Fonga, “Pattern Recognition in Gray-level Images by Fourier Analysis,” *Pattern Recognition Letters*, vol. 17, no. 14, pp. 1477–1489, 1996.
- [7] Mladen V. Wickerhauser, *Adapted Wavelet Analysis from Theory to Software*, IEEE Press, Piscataway, NJ, Estados Unidos, 1994.
- [8] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*, Addison-Wesley Publishing Company, Reading, Massachusetts, Estados Unidos, 1992.
- [9] Buh-Yun Lee, T.S. Liu, and Long-Wen Chang, “A New Optimal Digital Half-toning Technique Based on the Discrete Cosine Transform,” *Signal Processing*, vol. 80, pp. 2617–2621, 2000.
- [10] A. Gupta and K. R. Rao, “A Fast Recursive Algorithm for the Discrete Sine Transform,” *IEEE Acoustics, Speech, and Signal Processing*, vol. 38, no. 3, pp. 553–557, 1990.
- [11] O. K. Ersoy, “A Comparative Review of Real and Complex Fourier-related Transforms,” *Proceedings of the IEEE*, vol. 82, no. 3, pp. 429–447, 1994.
- [12] Ming C. Yang and Ja-Ling Wu, “A Novel Interpretation of the Two-Dimensional Discrete Hartley Transform,” *Signal Processing*, vol. 54, pp. 173–178, 1996.