

## 2 Transformações Integrais Invariáveis

Transformações bidimensionais tais como Fourier, Mellin e Transformada Z são de grande importância para o processamento de imagens [18, 20, 21]. O uso de tais transformadas pode ser verificado extensamente em aplicações como pré-processamento de imagens, filtragem e compressão.

No presente trabalho as Transformadas de Fourier, Mellin do tipo 1 e Mellin do tipo 2 são apresentadas e estudadas porque estas apresentam propriedades que permitem que imagens sejam transformadas em versões invariáveis à escala, distorções, translação e rotação. Estas versões permitem que imagens que sofreram as variações citadas possam ser diretamente comparadas umas com as outras.

É importante ressaltar que ao longo do trabalho o uso do termo distorção será referente a dilatações horizontais e verticais, ou seja, alterações em escala horizontais e verticais isoladamente. O termo não é usado para se referir a distorções locais.

Primeiramente é apresentada a Transformada de Fourier que possui magnitude invariável à translação.

Após a Transformada de Fourier, é apresentado o conceito da Transformada de Mellin de uma dimensão, transformada esta que possui características similares à Transformada de Fourier e possui magnitude invariável à escala.

Então, são apresentados dois métodos de mapear uma imagem:

- Mapeamento *Log-Log*: Possui a propriedade de transformar variações em escala horizontal e vertical em variações em translação;
- Mapeamento *Log-Polar*: Possui a propriedade de transformar variações em escala e rotação em variações em translação;

Após os mapeamentos serem explicados, são apresentados métodos de combinar a Transformada de Fourier e estes mapeamentos de modo a se construir as transformadas de Mellin do tipo 1 e 2 e de Fourier Mellin do tipo 1 e 2. Neste momento são apresentados meios de se chegar às invariâncias desejadas.

Por fim, é descrito como fazer a comparação entre imagens aproveitando-se das transformadas citadas e um método de *Window Growing* é proposto, método este que será utilizado durante a etapa de navegação para lugares conhecidos.

Este capítulo apresenta as seguintes seções:

- 2.1 – Transformada de Fourier: Apresenta a Transformada de Fourier para uma dimensão, para duas dimensões, suas versões discretas e suas características. É importante verificar a propriedade de translação que será explorada para se conseguir as invariâncias desejadas. A seção possui as seguintes subdivisões:
  - 2.1.1 – Transformada de Fourier Discreta;
  - 2.1.2 – Propriedade de Translação da DFT;
  - 2.1.3 – Propriedade de Escala da DFT;
  - 2.1.4 – Propriedade de Rotação da DFT;
- 2.2 – Transformada de Mellin: Descreve o conceito da Transformada de Mellin para funções de uma dimensão e mostra as características comuns desta com a Transformada de Fourier;
- 2.3 – Mapeamento *Log-Polar*: O mapeamento *Log-Polar* transforma uma imagem que está no sistema de coordenadas cartesiano para um novo sistema de coordenadas no qual variações em escala e rotação representam variações em translação. Esta característica será aproveitada junto à Transformada de Fourier para se chegar às transformações invariáveis à escala, translações e rotação;
- 2.4 – Mapeamento *Log-Log*: Este é outro mapeamento como o anterior. Este, porém, possui a característica de transformar escalas horizontais e verticais em translações;
- 2.5 – Obtendo Invariâncias à Rotação, Translação, Dilatações e Escala: Esta seção mostra como os conceitos apresentados até a mesma são combinados de modo a se obter as transformadas de Mellin do tipo 1 e 2, e Fourier Mellin do tipo 1 e 2. A seção está dividida em:
  - 2.5.1 – Invariância à Translação;
  - 2.5.2 – Invariância à Rotação e Escala;
  - 2.5.3 – Invariância a Dilatações Horizontais e Verticais;
  - 2.5.4 – Propriedade Comutativa Entre Rotação e Dilatação;

- 2.5.5 – Invariância à Translação, Rotação e Escala Simultâneas;
- 2.5.6 – Invariância à Translação e Dilatações Verticais e Horizontais Simultâneas;
- 2.6 – Comparando Imagens: A última seção do capítulo apresenta o uso das transformadas descritas para comparação entre imagens. As sub-seções desta são:
  - 2.6.1 – Distância Euclidiana;
  - 2.6.2 – Correlação;
  - 2.6.3 – Comparando Imagens com Invariâncias;
  - 0 – *Window Growing*;

## 2.1. Transformada de Fourier

Em uma dimensão, a Transformada de Fourier de uma função complexa  $f(x)$  é definida como:

$$F(k) = \mathfrak{T}[f(x)](k) = \int_{-\infty}^{\infty} f(x)e^{-2\pi i k x} dx \quad (1)$$

A Transformada de Fourier inversa é dada por:

$$f(x) = \mathfrak{T}^{-1}[F(k)](x) = \int_{-\infty}^{\infty} F(k)e^{2\pi i k x} dk \quad (2)$$

A versão em duas dimensões da Transformada de Fourier é dada por:

$$F(\omega_x, \omega_y) = \mathfrak{T}[f(x, y)](\omega_x, \omega_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)e^{-2\pi i (\omega_x x + \omega_y y)} dx dy \quad (3)$$

Sendo sua inversa:

$$f(x, y) = \mathfrak{T}^{-1}[F(\omega_x, \omega_y)](x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\omega_x, \omega_y)e^{2\pi i (k_x x + k_y y)} d\omega_x d\omega_y \quad (4)$$

Se a função  $f(x, y)$  apresenta intensidades ao longo das coordenadas espaciais  $x, y$ , então  $\omega_x, \omega_y$  são freqüências espaciais que representam as variações de intensidade em relação às distâncias espaciais.

Na próxima seção é apresentada a versão discreta da Transformada de Fourier.

### 2.1.1. Transformada de Fourier Discreta

Para uma função discreta de uma variável  $f(x_k)$  dada para  $x_k$  inteiro e definido no intervalo  $0 \leq x_k < N$ , sua Transformada de Fourier Discreta (DFT)  $F(u)$  é dada por:

$$F(u) = \sum_{x_k=0}^{N-1} f(x_k) e^{-2\pi i u x_k / M} \quad (5)$$

Sendo sua inversa:

$$f(x_k) = \frac{1}{N} \sum_{x_k=0}^{N-1} F(u) e^{2\pi i u x_k / M} \quad (6)$$

A DFT bidimensional de uma função discreta de duas variáveis  $f(x_k, y_j)$ , dada para  $x_k$  inteiro e definido no intervalo  $0 \leq x_k < N$  e  $y_k$  inteiro e definido no intervalo  $0 \leq y_k < M$ , é dada por:

$$F(u, v) = \sum_{y_k=0}^{M-1} \sum_{x_k=0}^{N-1} f(x_k, y_k) e^{-2\pi i (u x_k / N + v y_j / M)} \quad (7)$$

E sua inversa:

$$f(x_k, y_j) = \frac{1}{NM} \sum_{v=0}^{M-1} \sum_{u=0}^{N-1} F(u, v) e^{2\pi i (u x_k / N + v y_j / M)} \quad (8)$$

Uma imagem representada por uma função  $f(x_k, y_k)$  real geralmente possui DFT complexa. Pode-se então ter uma representação em amplitude e fase da DFT:

$$A(u, v) = |F(u, v)| \quad (9)$$

$$\Phi(u, v) = \angle F(u, v) \quad (10)$$

$$F(u, v) = |F(u, v)| e^{i \angle F(u, v)} \quad (11)$$

Serão vistas a seguir algumas propriedades da Transformada de Fourier Discreta. A primeira propriedade apresentada é em relação ao comportamento da Transformada de Fourier para funções transladadas.

### 2.1.2. Propriedade de translação da DFT

Translações espaciais na imagem em  $x_k$  ou  $y_k$  causam translações na fase da DFT:

$$f(x_k + a, y_k + b) \leftrightarrow F(u, v) e^{-i(au/N + bv/M)} \quad (12)$$

Perceba que tanto  $F(u, v)$  quanto  $f(x_k, y_j)$  são funções periódicas, portanto está implicitamente assumido que as translações são circulares.

É interessante verificar que a amplitude da DFT é invariável a translações na imagem. Isso quer dizer que duas imagens idênticas, porém transladadas, apresentariam a mesma amplitude após a Transformação de Fourier.

A próxima propriedade a ser vista é em relação a variações em escala na função a ser transformada. É interessante notar que variações em escala na função  $f$  são acompanhadas por variações semelhantes, porém inversas na magnitude de sua Transformada de Fourier.

### 2.1.3. Propriedade de Escala da DFT

Outra propriedade importante da Transformada de Fourier é que variações em escala no domínio espacial causam variações inversas no domínio da frequência:

$$f(\rho x_k, \alpha y_k) \leftrightarrow \frac{1}{\rho \alpha} F\left(\frac{u}{\rho}, \frac{v}{\alpha}\right) \quad (13)$$

Rotações na função  $f$  são acompanhadas por rotações na magnitude da Transformada de Fourier. Isto pode ser visto na seção seguinte.

### 2.1.4. Propriedade de Rotação da DFT

Rotações na imagem  $f(x_k, y_j)$  de um ângulo  $\theta$  fazem com que a sua Transformada de Fourier possua a mesma rotação  $\theta$ :

$$\begin{aligned} f(x_k \cos \theta - y_j \sin \theta, x_k \sin \theta + y_j \cos \theta) \leftrightarrow \\ F(u \cos \theta - v \sin \theta, u \sin \theta + v \cos \theta) \end{aligned} \quad (14)$$

Perceba que o *grid* é rotacionado, portanto o novo *grid* pode conter pontos não definidos. Os valores devem ser definidos então por interpolação.

Esta seção termina de apresentar a Transformada de Fourier. A seção seguinte apresentará a Transformada de Mellin e suas semelhanças com a Transformada de Fourier.

## 2.2. Transformada de Mellin

A Transformada de Mellin de uma função  $f(x)$  é definida em [23, 24] como:

$$M(s) = \int_0^{\infty} f(x)x^{s-1}dx \tag{15}$$

A partir desta definição pode-se perceber que a Transformada de Fourier de  $f(\exp \xi)$  é a Transformada de Mellin de  $M(i\omega)$  de  $f(x)$  ao longo do eixo imaginário do plano complexo. Esta relação pode ser vista através das seguintes considerações:

$$x = e^{\xi} \tag{16}$$

$$dx = -e^{\xi} d\xi \tag{17}$$

Aplicando-se a eq.(16) e a eq. (17) na eq. (15) temos:

$$M(s) = \int_0^{\infty} f(x)x^{s-1}dx = \int_{-\infty}^{\infty} f(e^{-\xi})e^{-s\xi}d\xi \tag{18}$$

Fazendo a substituição abaixo:

$$s = \theta + i^2\omega \tag{19}$$

Temos (tipicamente  $\theta = 0$ ):

$$M(i\omega) = \int_{-\infty}^{\infty} f(e^{-\xi})e^{-\sigma\xi}e^{-i2\pi\omega\xi}d\xi = \Im\left(f(e^{-\xi})\right)^{-\sigma\xi} \tag{20}$$

Esta transformada possui a propriedade de ter amplitude invariável à escala:

$$f(ax) \leftrightarrow M(i\omega)e^{-i(\omega \ln(a))} \tag{21}$$

Isto pode ser verificado como a seguir. Caso se escale a função  $f(x)$  por um fator  $\rho$ , o seguinte resultado é obtido:

$$\int_0^{\infty} f(\rho x) x^{i\omega-1} dx = \int_0^{\infty} f(x) \left(\frac{x}{\rho}\right)^{i\omega-1} \frac{dx}{\rho} = \rho^{-i\omega} \int_0^{\infty} f(x) x^{i\omega-1} dx = \rho^{-i\omega} M(i\omega) \quad (22)$$

Portanto, a Transformada de Mellin de uma função escalada adquire translação em sua fase porém mantém sua amplitude. Isto pode ser visto pelo fato de:

$$z^c = e^{c \ln z} = e^{c(\ln|z| + i \arg z)} = e^{c \ln|z|} e^{ic \arg z} \quad (23)$$

Para  $z = \rho$  e  $c = -i\omega$ , tem-se  $\arg(\rho) = 0$  e, portanto:

$$\begin{aligned} \rho^{-i\omega} &= e^{-i\omega \ln|\rho|} e^0 = \cos(-\omega \ln|\rho|) + i \sin(-\omega \ln|\rho|) \\ \rho^{-i\omega} (x + iy) &= \rho^{-i\omega} (r e^{i\theta}) = e^{-i\omega \ln|\rho|} (r e^{i\theta}) = r e^{i(\theta - \omega \ln|\rho|)} \end{aligned} \quad (24)$$

Isto corresponde somente a uma translação na fase dado que  $|e^{i\theta}| = 1$ .

Outro modo de se verificar a invariância na amplitude da transformada é pensar na Transformada de Mellin como uma Transformada de Fourier de uma função  $f_l(\xi)$  onde:

$$f_l(\xi) = f(e^\xi) = f(x) \quad (25)$$

Ou seja  $f_l$  é uma função que mapeia  $f$  da coordenada  $(x)$  para a coordenada  $(\xi)$ :

$$\xi = \ln(x) \quad (26)$$

Desta maneira, uma variação em escala em  $(x)$  representa uma translação em  $(\xi)$ :

$$(\rho x) \leftrightarrow (\xi + \ln \rho) \quad (27)$$

Como pode ser verificado por:

$$\xi' = \ln(\rho x) = \ln(\rho) + \ln(x) = \xi + \ln(\rho) \quad (28)$$

Portanto, como a Transformada de Fourier é invariável em amplitude à translação, e à variação em escala em  $f(x)$  corresponde a uma translação em  $f_l(\xi)$ , pode-se concluir que a Transformada Fourier de  $f_l(\xi)$  tem amplitude invariável para escalas em  $f(x)$ .

As Transformadas de Mellin do tipo 1 e 2 para imagens em 2 dimensões utilizam de propriedades dos mapeamentos *Log-Polar* e *Log-Log* que são apresentados nas seções 2.3 e 2.4 a seguir.

### 2.3. Mapeamento *Log-Polar*

Os mapeamentos de imagens para coordenadas distintas das originais são utilizados quando se deseja utilizar características específicas das novas coordenadas.

Serão apresentados aqui dois mapeamentos que possuem características interessantes quando aplicados junto à Transformada de Fourier. O primeiro a ser apresentado é o mapeamento *Log-Polar* que converte variações em escala e rotação para translações. O segundo é o mapeamento *Log-Log* que converte dilatações horizontais e verticais em translações.

A representação *Log-Polar* de uma imagem é feita através da transformação das coordenadas  $(x,y) \in \mathcal{R}^2$  para as coordenadas  $(\mu, \theta)$  como definido em [18, 23] como a seguir:

$$x = e^\mu \cos \theta \tag{29}$$

$$y = e^\mu \text{sen} \theta \tag{30}$$

$$\mu = \ln \sqrt{x^2 + y^2} \tag{31}$$

$$\theta = \tan^{-1} \frac{y}{x} \tag{32}$$

Onde  $\mu \in \mathcal{R}$  e  $0 \leq \theta \leq N$ .

Aplicando-se a eq. (29) e a eq.(30) temos o mapeamento *Log-Polar*  $f_p(\mu, \theta)$  de  $f(x,y)$ :

$$f_p(\mu, \theta) = f(x,y) = f(e^\mu \cos \theta, e^\mu \text{sen} \theta) \tag{33}$$

Este novo sistema de coordenadas é interessante devido ao fato de que escala e rotação são convertidas para translações.

Escala é convertida em translação na coordenada  $\mu$ :

$$(\rho x, \rho y) \leftrightarrow (\mu + \ln \rho, \theta) \tag{34}$$

Como pode ser verificado abaixo:

$$x' = \rho x \tag{35}$$

$$y' = \rho y \tag{36}$$

$$\mu' = \ln \sqrt{\rho^2 x^2 + \rho^2 y^2} = \ln \rho \sqrt{x^2 + y^2} = \tag{37}$$

$$\ln \rho + \ln \sqrt{x^2 + y^2} = \ln \rho + \mu$$

$$\theta = \tan^{-1} \frac{\rho y}{\rho x} = \theta \quad (38)$$

Rotação é convertida para translação na coordenada  $\theta$ :

$$(x \cos(\delta) - y \sin(\delta), x \sin(\delta) + y \cos(\delta)) \leftrightarrow (\mu, \theta + \delta) \quad (39)$$

Como demonstrado a seguir:

$$x' = x \cos(\delta) - y \sin(\delta) \quad (40)$$

$$y' = x \sin(\delta) + y \cos(\delta) \quad (41)$$

$$\begin{aligned} \mu' &= \ln \sqrt{(x \cos(\delta) - y \sin(\delta))^2 + (x \sin(\delta) + y \cos(\delta))^2} = \\ &= \ln \sqrt{x^2 \cos^2(\delta) + x^2 \sin^2(\delta) + y^2 \cos^2(\delta) + y^2 \sin^2(\delta) + 2xy \sin(\delta) \cos(\delta) - 2xy \sin(\delta) \cos(\delta)} = \\ &= \ln \sqrt{x^2 + y^2} = \mu \end{aligned} \quad (42)$$

$$\begin{aligned} \theta' &= \tan^{-1} \frac{x \sin \delta + y \cos \delta}{x \cos \delta - y \sin \delta} = \\ &= \tan^{-1} \frac{x \sin \delta + x \tan \theta \cos \delta}{x \cos \delta - x \tan \theta \sin \delta} = \\ &= \tan^{-1} \frac{\cos \theta \sin \delta + \sin \theta \cos \delta}{\cos \theta \cos \delta - \sin \theta \sin \delta} = \\ &= \tan^{-1} \frac{\sin(\theta + \delta)}{\cos(\theta + \delta)} = \tan^{-1} \tan(\theta + \delta) = \theta + \delta \end{aligned} \quad (43)$$

Estas conversões de escala e rotação para translação serão utilizadas junto à Transformada de Fourier para se obter transformadas invariáveis a estas características na seção 2.5. A seção seguinte apresenta outro mapeamento parecido com o descrito, mas que transforma variações em escala e rotação para variações em translação.

## 2.4. Mapeamento Log-Log

De modo similar ao mapeamento *Log-Polar*, pode-se fazer a transformação das coordenadas  $(x, y) \in \mathcal{R}^2$  para as coordenadas  $(\xi, \eta)$  como definido:

$$x = e^\xi \quad (44)$$

$$y = e^\eta \quad (45)$$

$$\xi = \ln x \quad (46)$$

$$\eta = \ln y \quad (47)$$

Onde  $\xi \in \mathcal{R}^2$  e  $\eta \in \mathcal{R}^2$ .

Desta vez ao aplicarmos a eq. (44) e a eq. (45) encontramos o mapeamento *Log-Log*  $f_i(\xi, \eta)$  de  $f(x, y)$ :

$$f_p(\xi, \eta) = f(x, y) = f(e^\xi, e^\eta) \quad (48)$$

Este mapeamento também possui propriedade interessante. Escalas independentes horizontais e verticais são traduzidas em respectivas translações:

$$(\rho x, \alpha y) \leftrightarrow (\xi + \ln \rho, \eta + \ln \alpha) \quad (49)$$

Como pode ser demonstrado:

$$x' = \rho x \quad (50)$$

$$y' = \alpha y \quad (51)$$

$$\xi' = \ln \rho x = \ln \rho + \ln x = \ln \rho + \xi \quad (52)$$

$$\eta' = \ln \alpha y = \ln \alpha + \ln y = \ln \alpha + \eta \quad (53)$$

As propriedades apresentadas nesta seção e na seção anterior serão utilizadas na seção seguinte para se conseguir obter transformadas baseadas na Transformada de Fourier que possuam invariância à escala, rotação e dilatações além da invariância à translação característica da própria Transformada de Fourier.

## 2.5. Obtendo Invariâncias à Rotação, Translação, Dilatações e Escala

Na presente seção os conceitos apresentados nas seções 2.1, 2.2, 2.3 e 2.4 serão utilizados para se definir as transformadas com invariâncias desejadas que serão aplicadas as imagens de modo a se poder fazer comparações entre imagens que representam uma mesma visão, mas com escalas, rotações, distorções e posições diferentes.

### 2.5.1. Invariância à Translação

Como foi apresentado na seção 2.1.4, a amplitude da Transformada de Fourier de uma imagem é invariável a translações espaciais circulares.

Veremos nas seções 2.5.2 e 2.5.3 que é possível aplicar esta invariância à translação de modo a cobrir mudanças em rotação, escala e escalas horizontais e verticais. Para isto será necessário trabalhar com os mapeamentos apresentados nas seções 2.3 e 2.4.

### 2.5.2. Invariância à Rotação e Escala

É possível se aplicar uma Transformada de Fourier de maneira peculiar para conseguir se obter invariância à escala e rotação.

Caso seja aplicada a Transformada de Fourier em uma imagem mapeada para o sistema de coordenadas *Log-Polar*, tem-se:

$$F(\varpi_x, \varpi_y) = \int_{-\infty}^{\infty} \int_0^{2\pi} f_p(\mu, \theta) e^{-2\pi i(\varpi_x \mu + \varpi_y \theta)} d\mu d\theta = \int_{-\infty}^{\infty} \int_0^{2\pi} f(e^\mu \cos \theta, e^\mu \sin \theta) e^{-2\pi i(\varpi_x \mu + \varpi_y \theta)} d\mu d\theta \tag{54}$$

A Transformada de Fourier da representação *Log-Polar* de uma imagem é conhecida como Fourier-Mellin [23]. A Transformada de Fourier-Mellin 2D é dada por:

$$M_2(\varpi_1, \varpi_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \frac{e^{-2\pi i(\varpi_1 \ln(x^2 + y^2) + \varpi_2 \tan^{-1} \frac{y}{x})}}{x^2 + y^2} dx dy \tag{55}$$

Uma translação na coordenada  $\mu$  representa uma mudança em tamanho no plano  $x$ - $y$ , e uma translação na coordenada  $\theta$  significa uma rotação no plano  $x$ - $y$ . Portanto, como a amplitude da Transformada de Fourier é invariável à translação para as duas variáveis, ao aplicarmos a mesma sobre  $f_p(\mu, \theta)$ , conseguimos na verdade invariância quanto à escala e rotação no plano  $x$ - $y$ . Doravante, escala e rotação em  $f(x, y)$  representam translações na fase da Transformada de Fourier Mellin:

$$f(\rho(x \cos \delta - y \sin \delta), \rho(x \sin \delta + y \cos \delta)) \leftrightarrow M(\varpi_1, \varpi_2) e^{-i(\varpi_1 \ln \rho + \varpi_2 \delta)} \tag{56}$$

É importante observar que ao se ganhar invariância à rotação e escala, se perde a invariância à translação em  $x$  e  $y$ .

A Transformada de Mellin apresentada é conhecida como Transformada de Mellin do tipo 2. A Transformada de Mellin do tipo 1 será apresentada na seção 2.5.3.

A Transformada de Mellin do tipo 2 pode ser discretizada levando-se em conta algumas considerações.

A função contínua  $f(x,y)$  pode ser discretizado em uma matriz  $F$  de tamanho  $N \times M$  cujos elementos são dados por  $f_{k,j}$  para uma imagem com *pixels* de dimensões  $\Delta x$  e  $\Delta y$ :

$$f(x,y) \rightarrow f(x_k,y_j) \text{ quando } f_{k,j} \equiv f(x_k,y_j) \quad (57)$$

$$x_k = k\Delta x \quad (58)$$

$$y_j = j\Delta y \quad (59)$$

$$f_{k,j} \equiv f(k\Delta x, j\Delta y) \quad (60)$$

Com  $k = 1, \dots, N$  e  $j = 1, \dots, M$ . A transformada contínua  $M_2(\omega_1, \omega_2)$  pode ser discretizada como uma matriz de tamanho  $U \times V$  considerando-se intervalos de frequência  $\Delta u/2\pi$  e  $\Delta v/2\pi$  para  $\omega_1$  e  $\omega_2$  respectivamente:

$$\omega_1 = \frac{u\Delta u}{2\pi} \quad (61)$$

$$\omega_2 = \frac{v\Delta v}{2\pi} \quad (62)$$

$$M_{(2)u,v} = M_2[\{f_{k,j}\}_{k=1}^N \{j=1}^M](u,v) \quad (63)$$

$$M_{(2)u,v} \equiv M_2(u \cdot \Delta u/2\pi, v \cdot \Delta v/2\pi) \quad (64)$$

Resultando em:

$$M_{(2)u,v} = \sum_{j=1}^M \sum_{k=1}^N \frac{1}{(k\Delta x)^2 + (j\Delta y)^2} f_{k,j} e^{iu\Delta u \ln((k\Delta x)^2 + (j\Delta y)^2)} e^{iv\Delta v \tan^{-1}\left(\frac{j\Delta y}{k\Delta x}\right)} \quad (65)$$

A seção seguinte apresentará a Transformada de Mellin do tipo 1 e como esta possui invariância a dilatações horizontais e verticais.

### 2.5.3. Invariância a Dilatações Horizontais e Verticais

De modo análogo ao apresentado na seção 2.5.2, pode-se conseguir invariância a dilatações horizontais e verticais.

Se a Transformada de Fourier for aplicada em uma imagem mapeada para o sistema de coordenadas *Log-Log*, obtém-se:

$$F(\varpi_x, \varpi_y) = \int_{-\infty}^{\infty} \int_0^{2\pi} f_l(\xi, \eta) e^{-2\pi i(\varpi_x \xi + \varpi_y \eta)} d\xi d\eta = \int_{-\infty}^{\infty} \int_0^{2\pi} f(e^\xi, e^\eta) e^{-2\pi i(\varpi_x \xi + \varpi_y \eta)} d\xi d\eta \quad (66)$$

Ao se considerar:

$$d\xi = -\frac{1}{x} dx \quad (67)$$

$$d\eta = -\frac{1}{y} dy \quad (68)$$

Esta transformada pode ser reescrita como a seguir, sendo conhecida como Transformada de Mellin do tipo 1:

$$M_1(\omega_1, \omega_2) = \int_0^{\infty} \int_0^{\infty} f(x, y) x^{i2\pi\omega_1-1} y^{i2\pi\omega_2-1} dx dy \quad (69)$$

$$M_1(\omega_1, \omega_2) = \int_0^{\infty} \int_0^{\infty} \frac{1}{xy} f(x, y) e^{i2\pi(\omega_1 \ln x + \omega_2 \ln y)} dx dy \quad (70)$$

Desta vez, uma dilatação no eixo  $x$  é traduzida para uma translação em  $\xi$  e uma dilatação no eixo  $y$  em uma translação em  $\eta$ . Portanto, a Fourier de  $f_l(\xi, \eta)$  acaba sendo invariável em amplitude a dilatações em  $x$  e  $y$ , estas representam translações na fase da Transformada de Mellin:

$$f(\rho x, \alpha y) \leftrightarrow M_1(\varpi_1, \varpi_2) e^{-i(\varpi_1 \ln \rho + \varpi_2 \ln \alpha)} \quad (71)$$

Como fazer uma variação em escala na função  $f(x, y)$  significa igualar as dilatações  $\rho$  e  $\eta$ , pode-se dizer que a amplitude da Transformada de Mellin do tipo 1 é invariável à escala:

$$f(\rho x, \rho y) \leftrightarrow M_1(\varpi_1, \varpi_2) e^{-i(\varpi_1 \ln \rho + \varpi_2 \ln \rho)} \quad (72)$$

Para se discretizar 0 basta aplicar as considerações vistas da eq. (57) à eq. (64) apresentadas obtendo-se:

$$M_{(1)u,v} = \sum_{j=1}^M \sum_{k=1}^N \frac{1}{k\Delta x} e^{iu\Delta u \ln(k\Delta x)} f_{k,j} \frac{1}{j\Delta y} e^{iv\Delta v \ln(j\Delta y)} \quad (73)$$

A Transformada de Mellin 2D do tipo 1 pode ser escrita de forma matricial como se segue:

$$Tx_{u,k} = \frac{1}{k} e^{iu\Delta u \ln(k\Delta x)} \tag{74}$$

$$Ty_{j,v} = \frac{1}{j} e^{iv\Delta v \ln(j\Delta y)} \tag{75}$$

$$M_1 = Tx \cdot F \cdot Ty \tag{76}$$

Onde:

$$T_x = \frac{1}{\Delta x} \begin{bmatrix} e^{i\Delta u \ln \Delta x} & \frac{1}{2} e^{i\Delta u \ln 2\Delta x} & \dots & \frac{1}{N} e^{i\Delta u \ln N\Delta x} \\ e^{i2\Delta u \ln \Delta x} & \frac{1}{2} e^{i2\Delta u \ln 2\Delta x} & \dots & \frac{1}{N} e^{i2\Delta u \ln N\Delta x} \\ \dots & \dots & \dots & \dots \\ e^{iU\Delta u \ln \Delta x} & \frac{1}{2} e^{iU\Delta u \ln 2\Delta x} & \dots & \frac{1}{N} e^{iU\Delta u \ln N\Delta x} \end{bmatrix} \tag{77}$$

$$T_y = \frac{1}{\Delta y} \begin{bmatrix} e^{i\Delta v \ln \Delta y} & e^{i2\Delta v \ln \Delta y} & \dots & e^{iV\Delta v \ln \Delta y} \\ \frac{1}{2} e^{i\Delta v \ln 2\Delta y} & \frac{1}{2} e^{i2\Delta v \ln 2\Delta y} & \dots & \frac{1}{2} e^{iV\Delta v \ln 2\Delta y} \\ \dots & \dots & \dots & \dots \\ \frac{1}{M} e^{i\Delta v \ln M\Delta y} & \frac{1}{M} e^{i2\Delta v \ln M\Delta y} & \dots & \frac{1}{M} e^{iV\Delta v \ln M\Delta y} \end{bmatrix} \tag{78}$$

A próxima seção explica a propriedade comutativa entre rotação e dilatação.

### 2.5.4. Propriedade Comutativa Entre Rotação e Dilatação

É interessante apontar que transformações em rotação dadas por  $R(\theta)$  e transformações em dilatação dadas por  $S(\rho, \alpha)$  comutam:

$$\begin{aligned} R(\theta)S(\rho, \alpha)f(x, y) &= R(\theta)f(\rho x, \alpha y) \\ &= f(\rho x \cos \theta - \rho y \sin \theta, \alpha x \sin \theta + \alpha y \cos \theta) \\ &= S(\rho, \alpha)f(x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta) \\ &= S(\rho, \alpha)R(\theta)f(x, y) \end{aligned} \tag{79}$$

Do mesmo modo transformações em escala dadas por  $S(\rho)$  poderiam ser usadas no lugar de  $S(\rho, \alpha)$  para mostrar que  $S(\rho)$  e  $R(\theta)$  comutam.

Similarmente, pode-se mostrar que transformações em translação  $T(\kappa, \tau)$  não comutam nem com  $S(\rho)$   $S(\rho, \alpha)$  nem  $R(\theta)$ .

Invariâncias à translação, rotação e escala por uma mesma transformada serão apresentadas a seguir utilizando a propriedade comutativa aqui mostrada.

### 2.5.5. Invariância à Translação, Rotação e Escala Simultâneas

Considere os dois operadores invariáveis  $F$  e  $F_{M2}$  apresentados a seguir:

$$F \circ f(x, y) = |F[f(x, y)](u, v)| \quad (80)$$

$$F_{M2} \circ f(x, y) = |M_2[f(x, y)](u, v)| \quad (81)$$

Como pode ser visto,  $F$ , extrai o modulo da Transformada de Fourier e  $F_{M2}$  extrai o módulo da Transformada de Mellin do tipo 2.

Aplicando-se o operador híbrido  $F \circ F_{M2}$  a uma imagem  $f(x, y)$  obtém-se:

$$I_1 = [F_{M2} \circ F]f(x, y) \quad (82)$$

Ao se aplicar o mesmo operador para uma imagem que foi transladada, rotacionada e escalada tem-se:

$$I_2 = [F_{m2} \circ F \circ R(\theta)S(\rho) \circ T(\kappa, \tau)]f(x, y) \quad (83)$$

$$= \left[ F_{M2} \circ R(\theta) \frac{1}{\rho^2} S\left(\frac{1}{\rho}\right) \circ F \circ T(\kappa, \tau) \right] f(x, y) \quad (84)$$

$$= \frac{1}{\rho^2} [F_{M2} \circ F]f(x, y) \quad (85)$$

$$= \frac{1}{\rho^2} I_1 \quad (86)$$

Portanto  $I_2$  é igual a  $I_1$  multiplicado por um fator  $1/\rho^2$  e a representação é invariável à escala, translação e rotação. Os passos 0. (83) e na eq. (84) se devem às propriedades apresentadas nas seções 2.1.3 e 2.1.4. A contração apresentada na eq. (85) se deve a propriedade de invariância de  $F$  e  $F_{M2}$ .

As invariâncias aqui descritas são suficientes para lidar com qualquer combinação ou permutação quanto à rotação, escala e translação em qualquer ordem.

De modo semelhante ao aqui apresentado, será apresentado adiante como se obter invariância à translação e dilatações verticais e horizontais.

### 2.5.6. Invariância à Translação e Dilatações Verticais e Horizontais Simultâneas

Considere o operador  $F$  apresentado na seção 2.5.5 e o operador  $F_{M1}$  a seguir:

$$F_{M1} \circ f(x, y) = |M_1[f(x, y)](u, v) \quad (87)$$

Aqui  $F_{M1}$  extrai o módulo da Transformada de Mellin do tipo 1 e  $F$  o módulo da Transformada de Fourier.

Aplicando-se o operador híbrido  $F \circ F_{M1}$  a uma imagem  $f(x, y)$  obtém-se:

$$I_3 = [F_{M1} \circ F]f(x, y) \quad (88)$$

Ao se aplicar o mesmo operador para uma imagem que foi transladada e dilatada nos eixos  $x$  e  $y$ , tem-se:

$$I_4 = [F_{m1} \circ F \circ R(\theta)S(\rho, \alpha) \circ T(\kappa, \tau)]f(x, y) \quad (89)$$

$$= \left[ F_{M1} \circ R(\theta) \frac{1}{\rho\alpha} S\left(\frac{1}{\rho\alpha}\right) \circ F \circ T(\kappa, \tau) \right] f(x, y) \quad (90)$$

$$= \frac{1}{\rho\alpha} [F_{M1} \circ F]f(x, y) \quad (91)$$

$$= \frac{1}{\rho\alpha} I_4 \quad (92)$$

Novamente encontra-se  $I_3$  proporcional a  $I_4$  e agora a representação é invariável às dilatações horizontal e vertical e à translação.

As transformadas invariáveis desenvolvidas serão utilizadas adiante para se comparar imagens.

### 2.6. Comparando Imagens

Todos os métodos apresentados para se encontrar funções invariáveis para imagens tem como fim a identificação e classificação das imagens. Para tal deseja-se comparar tais funções para diferentes imagens de modo a se obter uma função de similaridade entre estas.

Aqui serão apresentadas duas medidas de similaridade para se fazer as comparações: distância euclidiana e correlação. Então, é explicado com usar estas funções para comparar as imagens.

Por fim, é proposto um método para se comparar imagens obtidas em diferentes distâncias para uma mesma cena chamado de *Window Growing*.

### 2.6.1. Distância Euclidiana

Uma medida padrão de similaridade entre duas funções  $f(x,y)$  e  $g(x,y)$  é dada pela distância Euclidiana entre estas:

$$\text{Error! Objects cannot be created from editing field codes.} \quad (93)$$

Quanto maior a distância  $d$ , menor a semelhança entre as funções.

Repare que esta função tem a desvantagem de ser extremamente sensível a grandes magnitudes para valores pontuais das funções. Além disso, esta distância não é normalizada para limites dados.

Outra medida de similaridade muito comum é a correlação, esta será descrita a seguir.

### 2.6.2. Correlação

A Correlação entre duas funções  $f(x,y)$  e  $g(x,y)$  é definida de sua maneira mais geral pela integral contínua:

$$C(x',y') = \iint f(p,q)g(p+x',q+y') \quad (94)$$

Sua forma discreta é dada por:

$$C(x',y') = \sum \sum f(p,q)g(p+x',q+y') \quad (95)$$

Sua versão normalizada é vista como se segue:

$$C(x',y') = \frac{\sum \sum f(p,q)g(p+x',q+y')}{\sqrt{\sum \sum f(p,q)^2 \sum \sum g(p,q)^2}} \quad (96)$$

Quando  $x' = 0$  e  $y' = 0$ , esta correlação é feita para as funções sobrepostas. É, conseqüentemente, uma medida de similaridade entre as funções.

A correlação normalizada pode ser entendida como uma multiplicação vetorial para vetores normalizados. Portanto, é uma boa medida de similaridade entre duas funções por não ser tão sensível a grandes magnitudes para valores pontuais como a distância Euclidiana.

A correlação normalizada é igual a 1 para funções idênticas, igual a 0 para funções completamente diferentes e igual a  $-1$  para funções opostas, ou seja, quando  $f = -g$ ;

O valor de máximo da eq. (96) indica  $x'$  e  $y'$  que apresentam a melhor sobreposição entre as funções. Este valor é invariável a translações.

Para se correlacionar duas imagens com tamanhos diferentes, a imagem a ser comparada pode ser redimensionada para o tamanho da imagem base.

O uso da distância euclidiana e da correlação para se comparar imagens tendo invariância à escala, dilatações, rotações e translações é apresentado adiante.

### **2.6.3. Comparando Imagens com Invariâncias:**

Aplicando-se a correlação ou à distância euclidiana nas diversas transformadas das imagens, é possível fazer comparações invariáveis a vários fatores. Verifique as possibilidades a seguir:

- Máximo da correlação entre duas imagens: Invariância à translação;
- Distância ou correlação entre magnitude da Transformada de Fourier de duas imagens: Invariância à translação;
- Distância ou correlação entre magnitude da Transformada de Mellin do tipo 1 de duas imagens: Invariância a distorções (dilatações);
- Distância ou correlação entre magnitude da Transformada de Mellin do tipo 2 de duas imagens: Invariância à escala e rotação;
- Distância ou correlação entre magnitude da Transformada de Mellin do tipo 1 da magnitude da Transformada de Fourier de duas imagens: Invariância a distorções e translação;
- Distância ou correlação entre magnitude da Transformada de Mellin do tipo 2 da magnitude da Transformada de Fourier de duas imagens: Invariância à translação, escala e rotação;

A seguir é explicado um método interessante para se fazer comparações que será utilizado para auxiliar a navegação de um robô para um lugar conhecido.

### 2.6.4 **Window Growing**

Em determinadas situações, deseja-se encontrar uma imagem em uma sub-janela de outra imagem. Este é o caso de um robô observando um mesmo fundo, com a mesma direção, porém para distâncias diferentes. Para tal é proposta uma técnica de “*window growing*” onde uma das imagens será sub-dividida em várias sub-imagens, para então, fazer a comparação.

Na Figura 2-1 ilustra como isto é feito. O retângulo à esquerda representa a imagem de referência que será procurada na outra imagem, representada pelo retângulo à direita. Perceba que a imagem à direita possui várias sub-divisões, ou melhor, sub-janelas. Estas sub-janelas serão comparadas com a imagem de referência.

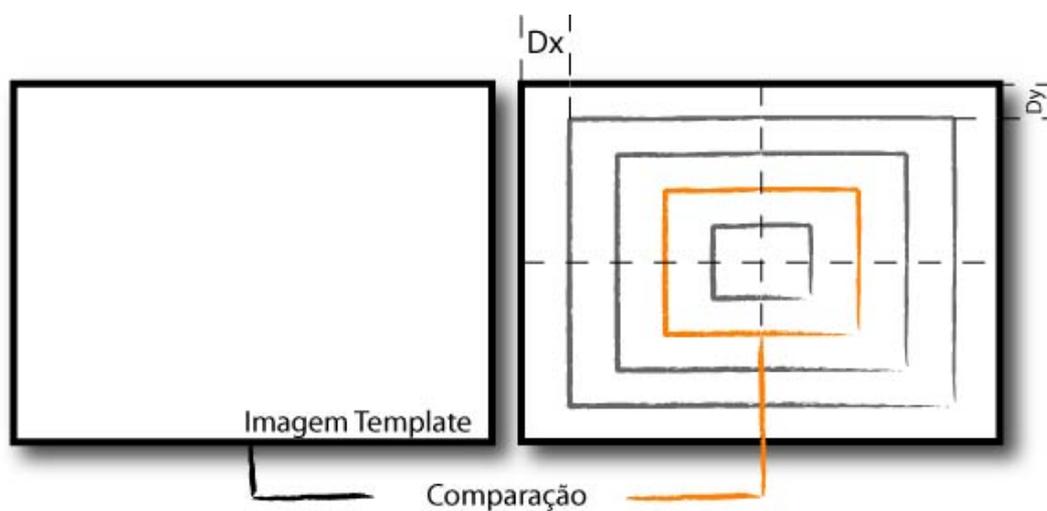


Figura 2-1: *Window Growing*

A figura na qual se deseja procurar a imagem base é subdividida em um número  $n$  de janelas com centro comum. As diferenças de tamanhos  $Dx$  e  $Dy$  entre uma janela e outra são dadas por:

$$Dx = sx/2n \tag{97}$$

$$Dy = sy/2n \tag{98}$$

Onde  $sx$  é o tamanho horizontal da imagem a ser comparada e  $sy$  o tamanho vertical.

Perceba que apesar das janelas a serem comparadas terem tamanhos diferentes, algumas transformadas apresentadas podem ser obtidas de modo que as

funções obtidas tenham tamanhos semelhantes. Isto pode ser feito para qualquer comparação que antes de ser feita passe por uma Transformada de Mellin do tipo 1 ou 2.

Para se comparar através de correlação ou Fourier, pode-se fazer o redimensionamento da imagem base para então fazer a comparação.

Por fim, tem-se uma série de comparações, uma para cada sub-janela e pode-se encontrar a janela com maior semelhança com a imagem base.

De modo a não ser necessário ter  $n$  muito grande para se encontrar uma sub-janela bem próxima da imagem base, pode-se fazer uma segunda tomada de comparações criando-se novas sub-imagens ao redor daquela que obteve melhor comparação.

Isto está exemplificado na Figura 2-2. Entre as duas sub-janelas ao redor daquela que obteve maior valor de comparação são sub-divididas novas  $m$  janelas e o processo é retomado.

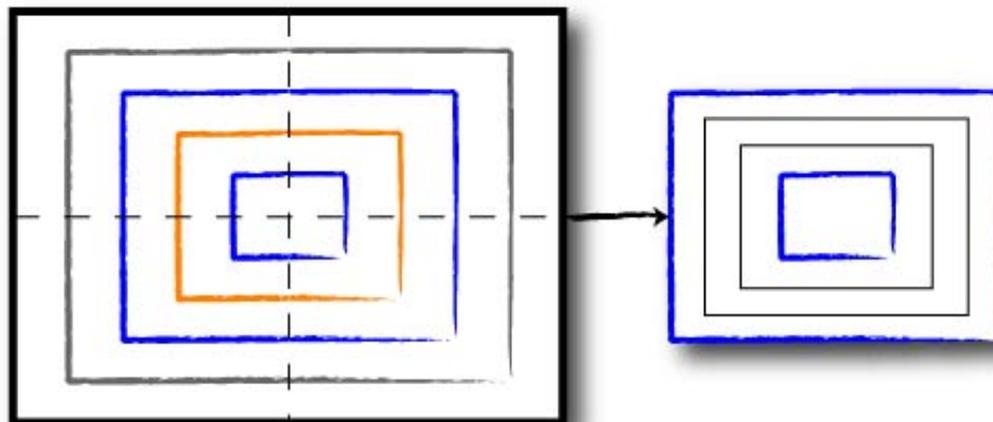


Figura 2-2: Nova tomada de comparações

Este método apresentado pode ser feito iterativamente o número de vezes necessário até que se chegue em um intervalo entre janelas desejado. O diagrama de fluxo do procedimento é dado na Figura 2-3.

O próximo capítulo apresentará uma técnica que será utilizada para se encontrar regiões locais em comum entre imagens.

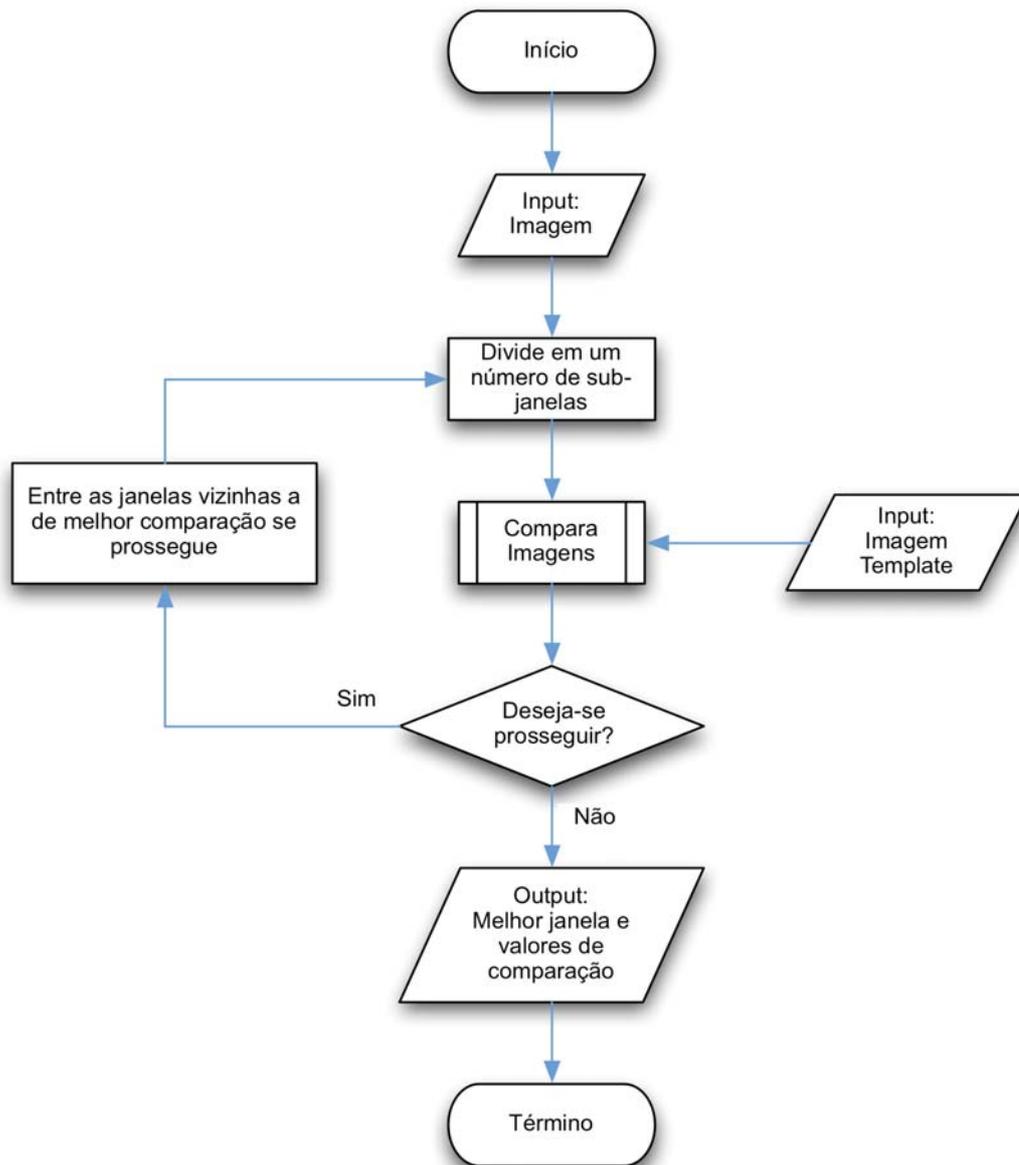


Figura 2-3: Diagrama de fluxo do procedimento de *Window Growing*