Filtering

- The image shown below has been generated by adding some noise in the form of a cosine function.

- If we denote the original image as \( f(x,y) \), then the noisy image can be denoted as \( f(x,y) + n(x,y) \) where \( n(x,y) \) is a cosine function.

- Using the additive property of FT, the FT of the noisy image will be \( F(u,v) + N(u,v) \) where \( F(u,v) \) is the FT of \( f(x,y) \) and \( N(u,v) \) is the FT of \( n(x,y) \).

- We can remove this noise effectively in the frequency domain using FT (filter out the frequencies corresponding to the cosine function)
Fourier Descriptors


• **Region Representation for 2D Recognition**

  - In terms of boundaries (external characteristics)
    
    (good when shape information is important)

  - In terms of pixels comprising the region (internal characteristics)
    
    (good when texture, color, etc. are important)

• **Region Description for 2D Recognition**

  - Extract a number of features (*descriptors*)
    
    (e.g., length, orientation of lines, concavities etc.)

  - Descriptors can be used for recognition

  - Must be invariant to size, translation, and rotation
• Fourier Descriptors

- Represent the boundary as a sequence of point coordinates by traversing the boundary, for example, in a counterclockwise order

\[(x_0, y_0), (x_1, y_1), \ldots, (x_{N-1}, y_{N-1})\]

(for simplicity, \(x(k) = x_k\) and \(y(k) = y_k\))

- Treat each coordinate as a complex number

\[z(k) = x(k) + jy(k), k = 0, 1, \ldots, N - 1\]

- Consider the FT of \(s(k)\)

\[\alpha(u) = \frac{1}{N} \sum_{k=0}^{N-1} z(k)e^{-j2\pi k/N}\]

- The complex coefficients \(\alpha(u)\) are called the *fourier descriptors* of the boundary
- Take the IFT

\[ z(k) = \sum_{u=0}^{N-1} \alpha(u)e^{j2\pi k/N} \]

- Represent the boundary using a few fourier descriptors

- Let’s consider the first \( M \) terms of the sum only

\[ \hat{z}(k) = \sum_{u=0}^{M-1} \alpha(u)e^{j2\pi k/N}, \ M < N \]
• **Properties**

- Fourier descriptors can be computed fast using FFT.

- The zero descriptor $\alpha(0)$ represents the center of gravity of the curve.

- A translation of the curve by $z_0 = x_0, y_0$ affects only the $\alpha(0)$ term:

  $$\alpha_t(0) = \alpha(0) + z_0$$

- A rotation of the curve by an angle $\theta$ results in a phase shift of the coefficients:

  $$\alpha_r(u) = \alpha(u)e^{i\theta}$$

- A scaling of the curve by a factor $s$ (about the center of gravity of the curve) results in scaling of the Fourier coefficients:

  $$z_s(k) = sz(k)$$
  $$\alpha_s(u) = s\alpha(u)$$

- A change in the starting point of curve traversal produces modulation of the Fourier descriptors:

  $$z_p(k) = z(k - k_0)$$
  $$\alpha_p(u) = \alpha(u)e^{-i2\pi k_0 u/N}$$

• **Normalization**

- Translation invariance: set $\alpha(0) = 0$

- Scale invariance: $\alpha'(n) = \frac{\alpha(n)}{|\alpha(1)|}$

  ($\alpha(1)$ will always be the largest assuming the curve does not cross itself)

- Rotation and starting point invariance: can be done but more complicated.
Image Watermarking


• Data hiding or steganography

- Steganography deals with methods of embedding data within a medium (e.g., images) in an imperceptible way.

- The invisibly embedded signal (watermark) completely characterizes the person who applied it on the image, and can be used as a proof of digital media ownership (copyright protection).

• Why it works?

- Data hiding techniques in images explore the noise or distortion masking property of the human visual system.

- This property refers to the decrease in the perceived intensity of a visual stimulus when this is superimposed over another stimulus.

• Related areas
Encryption

- Data transmitted over a network may be protected from unauthorized receivers using cryptography techniques.

- Only persons who possess the appropriate private key can decrypt the data using a public algorithm.

- Once the data is decrypted, however, it can be freely distributed or manipulated.

Authentication

- It is related to data integrity verification.

- Should be able to signal a data integrity violation when significant modification of the visual content occur and pinpoint the image regions which have been altered.

- In contrast to copyright protection, data inserted for authentication should be fragile (i.e., should be modified when the image is manipulated).

Covert communications

- Exchange of messages secretly embedded within images.

- Being able to convey a fairly large amount of data is a basic requirements for such applications.
• Requirements

- The modifications caused by watermark embedding should not degrade the perceived image quality.

- Watermark signals should be characterized by great complexity to be able to produce an extensive set of well distinguishable watermarks (trustworthy detection).

- Watermarks should be associated with a key which is used to cast, detect, or remove a watermark (private or public keys).

- The watermark must be easily and securely detectable by the owner (i.e., using the key).

- The watermark must be difficult or impossible to remove (i.e., without the key), at least without visibly degrading the original image.

- The watermark must survive image modifications that are common to typical applications (e.g., filtering, compression).

- Watermarks should not be recovered using statistical methods (i.e., should be image dependent).

• Steps in watermarking

  **Generation:** producing the watermark pattern, using an owner and/or image dependent key.

  **Embedding:** a superposition of the watermark signal on the original image.

  **Detection:** verifying whether the image under consideration hosts a certain watermark.
• Categorization of watermarking techniques

Availability of original image during detection

(1) Private or image escrow schemes

* Require that the original image is available during detection.
* Robust to a wide range of image manipulation including filtering and geometric manipulations such as cropping, scaling, rotation, etc.
* Do not fit well in certain applications.

(2) Oblivious or blind methods

* Do not require the original image for watermark detection.
* More general compared to private schemes.
* Limited robustness to image modifications and attacks.

Domain where the watermark signal is embedded

(1) Spatial domain methods

* Do data embedding in the spatial domain by modulating the intensity of certain pixels.

(2) Transform domain methods

* Modify the magnitude of the coefficients in an appropriate transform domain (e.g., DFT, DWT, DCT etc.)
Rotation, Scale and Translation Invariant
Digital Image Watermarking

• Overview of the method

- It is based on the Fourier-Mellin transform.

- The embedded marks are unaffected by any combination of rotation, scale, and translation.

- Watermark is encoded using spread spectrum encoding:
  * Encoded watermark resembles a pseudo-random sequence.
  * This allows embedding messages of arbitrary length.
  * A secret key is needed for decoding the watermark (seed used to generate the pseudo-sequence).

- The encoding scheme is coupled with an error-correcting scheme for improved robustness.
• Fourier-Mellin transform

- It’s the Fourier transform applied on a log-polar coordinate system.

- The log-polar transform is defined as follows:

\[
x = e^\mu \cos(\theta) \\
y = e^\mu \sin(\theta)
\]

where \( \mu \in R \) and \( 0 \leq \theta < 2\pi \)

- There is a unique correspondence between \((x, y)\) and \((\mu, \theta)\)
• **Rotation and scale invariance**

- Scaling in \((x, y)\) domain corresponds to translation in \((\mu, \theta)\) domain.

\[
(sx, sy) \rightarrow (\mu + \log(s), \theta)
\]

- Rotation in \((x, y)\) domain corresponds to translation in \((\mu, \theta)\) domain.

\[
(x \cos(\phi) - y \sin(\phi), x \sin(\phi) + y \cos(\phi)) \rightarrow (\mu, \theta + \phi)
\]

- The magnitude of the Fourier-Mellin transform is invariant to translation.

• **Rotation, scale and translation invariance**

- Consider the following two operators:

1. \(F\) which extracts the magnitude of the Fourier transform.
2. \(F_M\) which extracts the magnitude of the Fourier-Mellin transform.

- Consider an image \(f(x, y)\) which is translated, scaled, and rotated

\[
\hat{f}(x, y) = R \ S \ T \ f(x, y)
\]

- It can be shown that

\[
[F_M \ F]f(x, y) = [F_M \ F]\hat{f}(x, y)
\]

(true for any combination of transformations and in any order)
- There are many practical issues (see paper for details)

* Computing the log-polar or inverse log-polar requires interpolation.
* Both steps introduce some loss of information.

**Watermark embedding and extraction**

- The watermark is a 1D spread spectrum signal as discussed.

- Each element is added to some percentage of the largest components in the invariant domain (i.e., Fourier-Mellin).

- The variance of the spread spectrum signal indicates the strength of the watermark.

- The diagram below illustrates the extraction of the watermark.
• **Experiments**

  - The authors demonstrate the robustness of their method to rotation, scale, translation.

  - Also, they demonstrate reasonable robustness to JPEG compression (50%-75% quality factor) and cropping (50%).

  - It takes 7 sec to embed a message and 5 sec to extract it.