

Rapport de Devoir Individuel

Descripteurs de Forme et de Texture

Module : Multimedia Mining and Indexing

Réalisé par :

Reda Wafik

Encadré par :

Prof. M'hamed AIT KBIR

Université Abdelmalek Essaâdi
Faculté des Sciences et Techniques de Tanger
Master IASD - Semestre 3

Année Universitaire 2025-2026

Table des matières

1	Introduction	3
2	Fondements Théoriques	3
2.1	Descripteurs de Forme	3
2.1.1	Descripteurs de Fourier	3
2.1.2	Histogramme des Directions	3
2.2	Descripteurs de Texture	4
2.2.1	Filtres de Gabor	4
2.2.2	Caractéristiques de Tamura	4
2.3	Mesure de Similarité	5
3	Méthodologie et Implémentation	5
3.1	Architecture du Système	5
3.2	Pipeline d'Extraction des Descripteurs de Forme	5
3.3	Pipeline d'Extraction des Descripteurs de Texture	6
3.4	Processus de Recherche d'Images Similaires	6
4	Résultats Expérimentaux	6
4.1	Résultats sur les Images de Formes	6
4.1.1	Requête : apple-2.gif	7
4.1.2	Requête : bell-3.gif	8
4.1.3	Requête : bird-7.gif	9
4.2	Résultats sur les Images de Textures	9
4.2.1	Requête : Im01.jpg	10
4.2.2	Requête : Im02.jpg	10
4.3	Analyse des Performances	11
4.3.1	Descripteurs de Forme	11
4.3.2	Descripteurs de Texture	11
5	Conclusion	11
5.1	Synthèse des Résultats	11
5.2	Choix Méthodologiques	11
5.3	Apports et Limites	12
5.4	Perspectives	12
5.5	Conclusion Générale	12

1 Introduction

L'indexation et la recherche d'images par le contenu (CBIR - Content-Based Image Retrieval) constituent des domaines fondamentaux du traitement d'images et de l'intelligence artificielle. Ce projet vise à développer un système de recherche d'images similaires basé sur deux types de descripteurs visuels : les descripteurs de forme et les descripteurs de texture.

L'objectif principal est d'extraire des caractéristiques pertinentes d'images, de les représenter sous forme de vecteurs numériques, puis de calculer des mesures de similarité permettant de retrouver les images les plus proches d'une image requête donnée.

Ce rapport présente la méthodologie suivie, les fondements théoriques des descripteurs utilisés, l'implémentation pratique avec Python et OpenCV, ainsi que les résultats obtenus sur deux ensembles de données distincts : un ensemble d'images de formes et un ensemble de textures.

2 Fondements Théoriques

2.1 Descripteurs de Forme

Les descripteurs de forme permettent de caractériser la géométrie et la silhouette d'objets dans une image. Nous avons utilisé deux types de descripteurs complémentaires :

2.1.1 Descripteurs de Fourier

Les descripteurs de Fourier sont basés sur la transformée de Fourier discrète (DFT) appliquée au contour d'un objet. Le contour est représenté comme une séquence de points complexes :

$$z(k) = x(k) + jy(k), \quad k = 0, 1, \dots, N - 1 \quad (1)$$

où $(x(k), y(k))$ sont les coordonnées du k -ième point du contour.

La transformée de Fourier discrète est donnée par :

$$F(u) = \frac{1}{N} \sum_{k=0}^{N-1} z(k) e^{-j2\pi uk/N}, \quad u = 0, 1, \dots, N - 1 \quad (2)$$

Les coefficients de Fourier $F(u)$ capturent les caractéristiques fréquentielles du contour. Les basses fréquences représentent la forme globale, tandis que les hautes fréquences codent les détails fins.

Pour obtenir une invariance à la translation, rotation et mise à l'échelle, nous normalisons les coefficients :

$$FD(u) = \frac{|F(u)|}{|F(1)|}, \quad u = 1, 2, \dots, n \quad (3)$$

Dans notre implémentation, nous conservons les 32 premiers coefficients normalisés.

2.1.2 Histogramme des Directions

L'histogramme des directions caractérise l'orientation locale du contour. Pour chaque segment du contour, nous calculons l'angle :

$$\theta_i = \arctan \left(\frac{\Delta y_i}{\Delta x_i} \right) \quad (4)$$

où $\Delta x_i = x_{i+1} - x_i$ et $\Delta y_i = y_{i+1} - y_i$.

Les angles sont ensuite quantifiés en 36 bins couvrant $[0, 360]$, produisant un histogramme normalisé :

$$H(b) = \frac{n_b}{\sum_{b'} n_{b'}}, \quad b = 0, 1, \dots, 35 \quad (5)$$

où n_b est le nombre d'angles dans le bin b .

2.2 Descripteurs de Texture

Les descripteurs de texture capturent les variations spatiales d'intensité dans une image.

2.2.1 Filtres de Gabor

Les filtres de Gabor sont des filtres linéaires utilisés pour l'analyse de texture. Un filtre de Gabor 2D est défini par :

$$g(x, y; \lambda, \theta, \sigma, \gamma) = \exp \left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2} \right) \cos \left(2\pi \frac{x'}{\lambda} \right) \quad (6)$$

où :

- $x' = x \cos \theta + y \sin \theta$
- $y' = -x \sin \theta + y \cos \theta$
- λ : longueur d'onde (pixels par cycle)
- θ : orientation du filtre
- σ : écart-type de l'enveloppe gaussienne
- γ : rapport d'aspect spatial

Nous appliquons une banque de filtres avec différentes orientations $\theta \in \{0, 45, 90, 135\}$, échelles $\sigma \in \{3, 5\}$ et longueurs d'onde $\lambda \in \{10, 20\}$.

Pour chaque réponse filtrée $R_i(x, y)$, nous extrayons :

$$\mu_i = \mathbb{E}[R_i(x, y)] \quad (7)$$

$$\sigma_i^2 = \text{Var}[R_i(x, y)] \quad (8)$$

Cela produit un vecteur de $2 \times 2 \times 4 \times 2 = 32$ caractéristiques.

2.2.2 Caractéristiques de Tamura

Les caractéristiques de Tamura modélisent la perception visuelle humaine de la texture :

Contraste : mesure la distribution des niveaux de gris

$$\text{Contraste} = \sigma_I \quad (9)$$

où σ_I est l'écart-type de l'image.

Directionnalité : quantifie l'orientation dominante

$$\text{Directionnalité} = \text{Var}[\nabla I] \quad (10)$$

où ∇I représente le gradient (obtenu par Sobel).

Rugosité (Roughness) : mesure les variations locales

$$\text{Rugosité} = \mathbb{E}[|\nabla^2 I|] \quad (11)$$

où $\nabla^2 I$ est le Laplacien de l'image.

Granularité (Coarseness) : évalue l'échelle spatiale dominante

$$\text{Granularité} = \mathbb{E}[\text{Blur}(I, k)] \quad (12)$$

où $\text{Blur}(I, k)$ est un flou moyen avec un noyau de taille $k \times k$.

2.3 Mesure de Similarité

Pour comparer deux images, nous calculons la distance euclidienne entre leurs vecteurs de caractéristiques :

$$d(\mathbf{v}_1, \mathbf{v}_2) = \|\mathbf{v}_1 - \mathbf{v}_2\|_2 = \sqrt{\sum_{i=1}^n (v_{1,i} - v_{2,i})^2} \quad (13)$$

Plus la distance est petite, plus les images sont similaires visuellement.

3 Méthodologie et Implémentation

3.1 Architecture du Système

Le système développé comprend quatre modules principaux :

1. **Module d'extraction des descripteurs de forme** : Traite les images GIF pour extraire les caractéristiques géométriques
2. **Module d'extraction des descripteurs de texture** : Traite les images JPG pour extraire les caractéristiques texturales
3. **Module de recherche de formes similaires** : Compare et retrouve les formes les plus proches
4. **Module de recherche de textures similaires** : Compare et retrouve les textures les plus proches

3.2 Pipeline d'Extraction des Descripteurs de Forme

Le processus d'extraction pour les images de formes suit ces étapes :

1. **Prétraitement** : Chargement de l'image en niveaux de gris
2. **Binarisation** : Application du seuillage automatique d'Otsu pour séparer l'objet du fond
3. **Détection des contours** : Utilisation de l'algorithme de détection de contours
4. **Sélection du contour principal** : Choix du contour avec le nombre maximum de points
5. **Calcul des descripteurs de Fourier** : Extraction de 32 coefficients normalisés
6. **Calcul de l'histogramme directionnel** : Quantification en 36 bins de 10° chacun
7. **Sauvegarde** : Stockage du vecteur caractéristique au format JSON

Le vecteur final pour chaque image de forme contient donc 68 dimensions (32 + 36).

3.3 Pipeline d'Extraction des Descripteurs de Texture

Pour les images de texture, le processus est le suivant :

1. **Prétraitement** : Chargement de l'image en niveaux de gris
2. **Banque de filtres de Gabor** : Application de 16 filtres différents
 - 2 échelles : $\sigma \in \{3, 5\}$
 - 4 orientations : $\theta \in \{0, 45, 90, 135\}$
 - 2 longueurs d'onde : $\lambda \in \{10, 20\}$ pixels
3. **Extraction des statistiques de Gabor** : Pour chaque réponse filtrée, calcul de la moyenne et de la variance
4. **Calcul des caractéristiques de Tamura** :
 - Contraste (écart-type global)
 - Directionnalité (variance du gradient)
 - Rugosité (moyenne du Laplacien en valeur absolue)
 - Granularité (moyenne du flou)
5. **Sauvegarde** : Stockage du vecteur caractéristique au format JSON

Le vecteur final pour chaque image de texture contient 36 dimensions (32 de Gabor + 4 de Tamura).

3.4 Processus de Recherche d'Images Similaires

L'algorithme de recherche implémente les étapes suivantes :

1. **Chargement** : Lecture du vecteur caractéristique de l'image requête depuis le fichier JSON
2. **Comparaison** : Calcul de la distance euclidienne entre le vecteur requête et tous les autres vecteurs de la base
3. **Classement** : Tri des résultats par distance croissante
4. **Sélection** : Conservation des 6 images les plus similaires
5. **Visualisation** : Création d'une grille 2×3 affichant l'image requête et les résultats avec leurs distances respectives
6. **Sauvegarde** : Enregistrement de la grille résultante au format PNG

Cette approche permet une recherche rapide et efficace dans des bases d'images de taille moyenne.

4 Résultats Expérimentaux

4.1 Résultats sur les Images de Formes

Nous avons testé le système sur trois images requêtes différentes. Les résultats montrent que le système identifie correctement les formes similaires en se basant sur leurs caractéristiques géométriques.

4.1.1 Requête : apple-2.gif

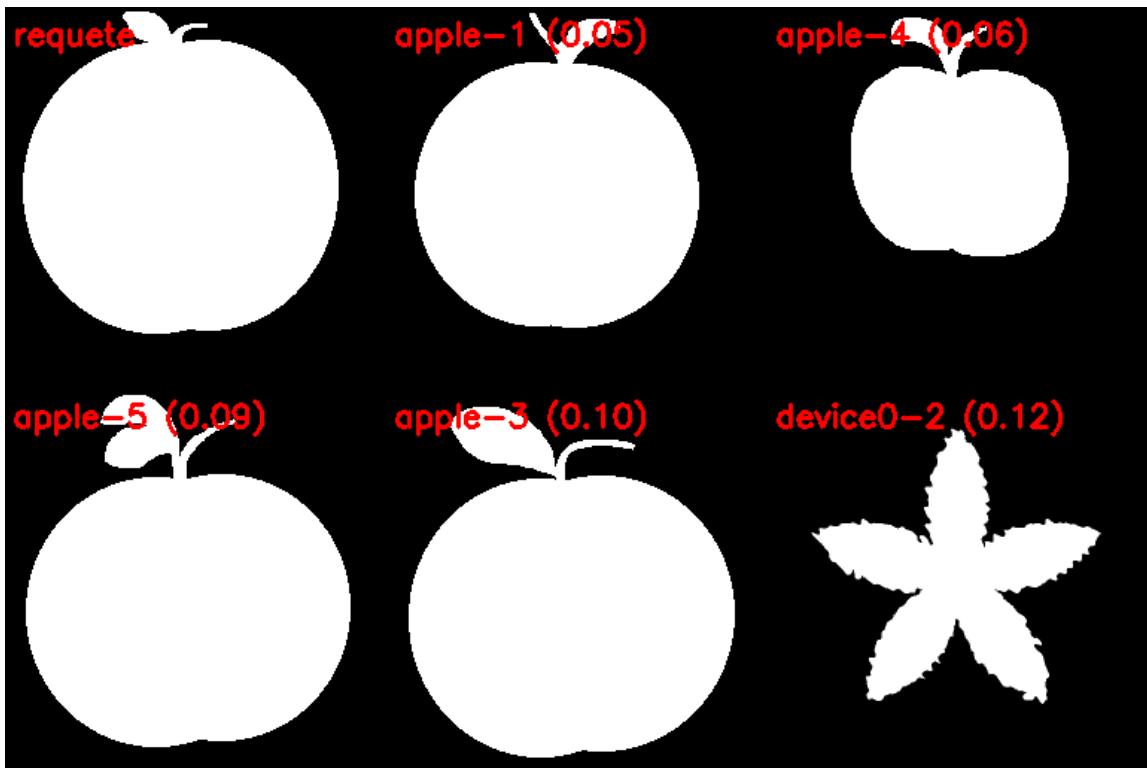


FIGURE 1 – Résultats de recherche pour l'image apple-2.gif

Les images retournées présentent des formes arrondies similaires à la pomme requête. Les distances calculées reflètent le degré de similarité géométrique.

4.1.2 Requête : bell-3.gif

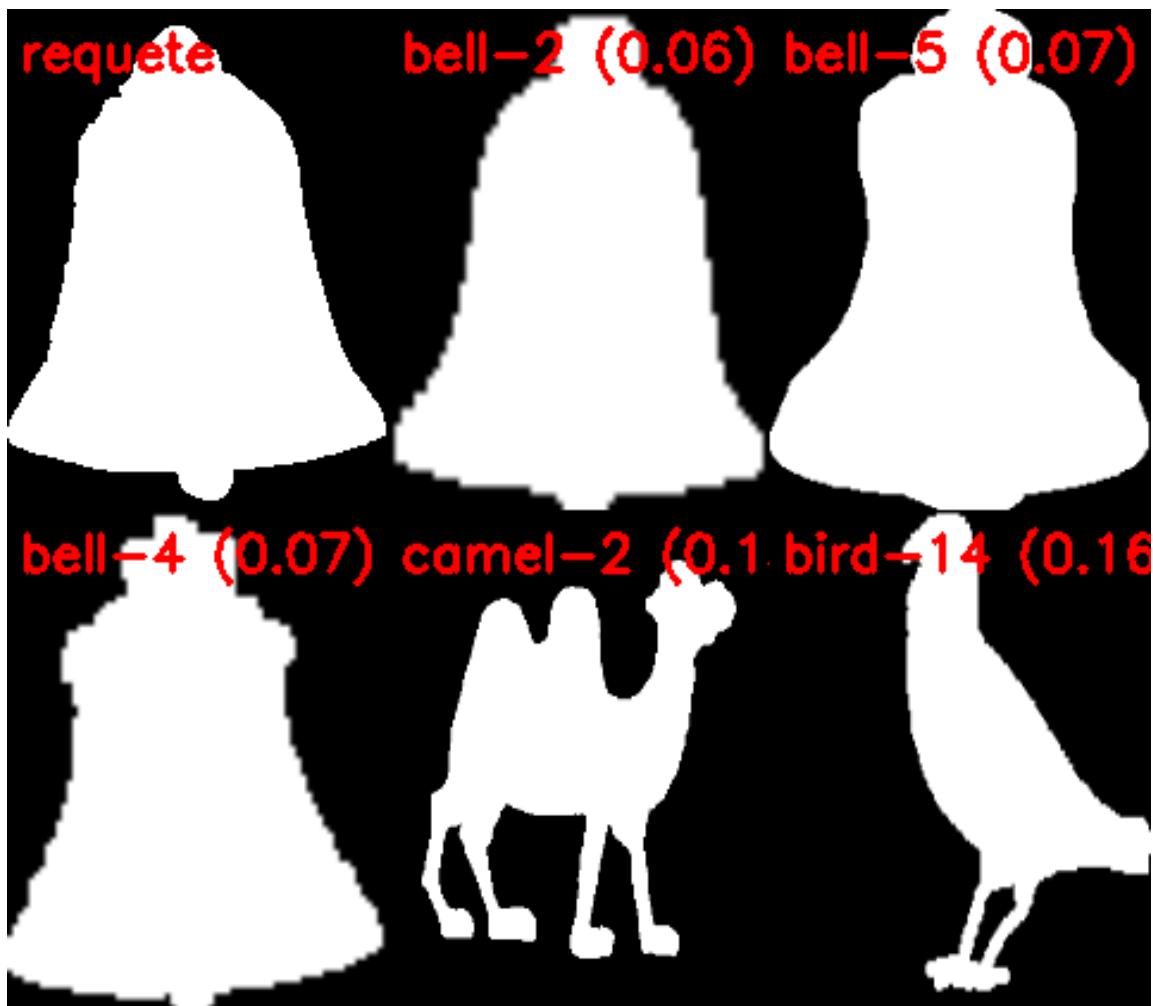


FIGURE 2 – Résultats de recherche pour l'image bell-3.gif

Le système identifie des formes avec des contours similaires à la cloche, notamment des objets présentant une base large et un sommet arrondi.

4.1.3 Requête : bird-7.gif

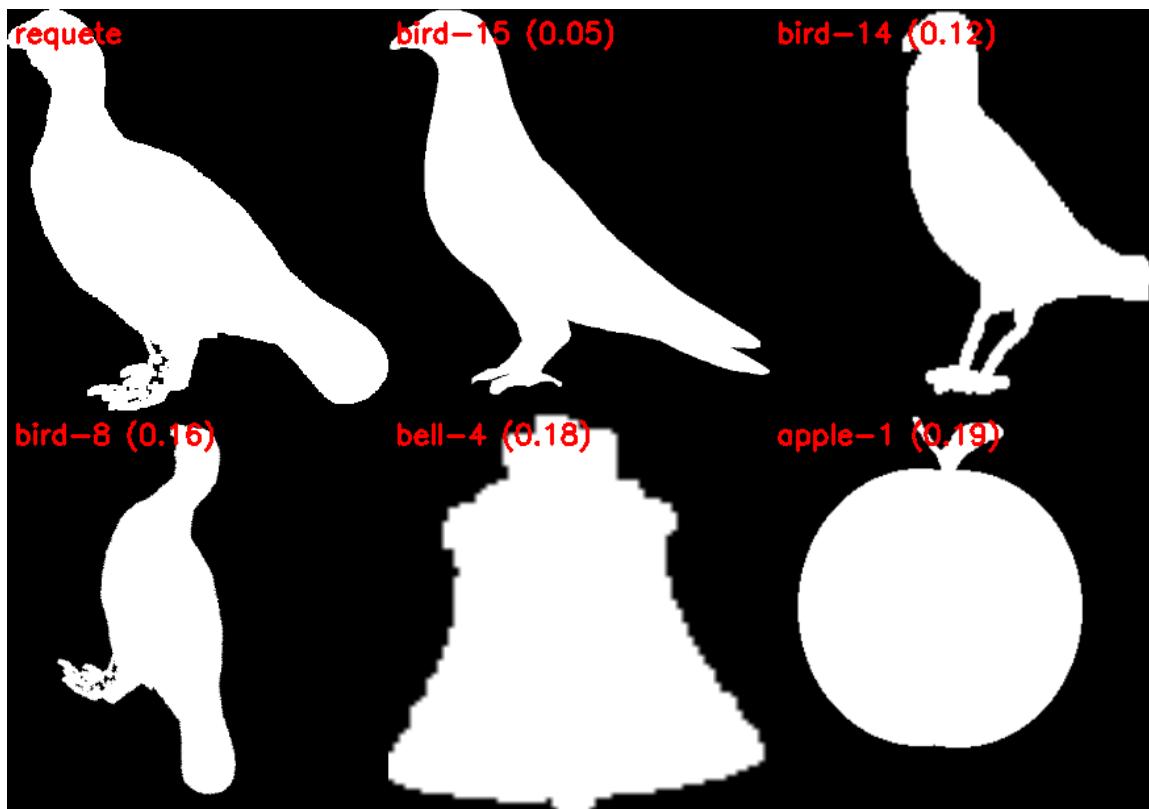


FIGURE 3 – Résultats de recherche pour l'image bird-7.gif

Les images retournées partagent des caractéristiques structurelles avec l'oiseau requête, notamment en termes de silhouette et de distribution directionnelle.

4.2 Résultats sur les Images de Textures

Les descripteurs de texture permettent d'identifier des images présentant des motifs visuels similaires.

4.2.1 Requête : Im01.jpg

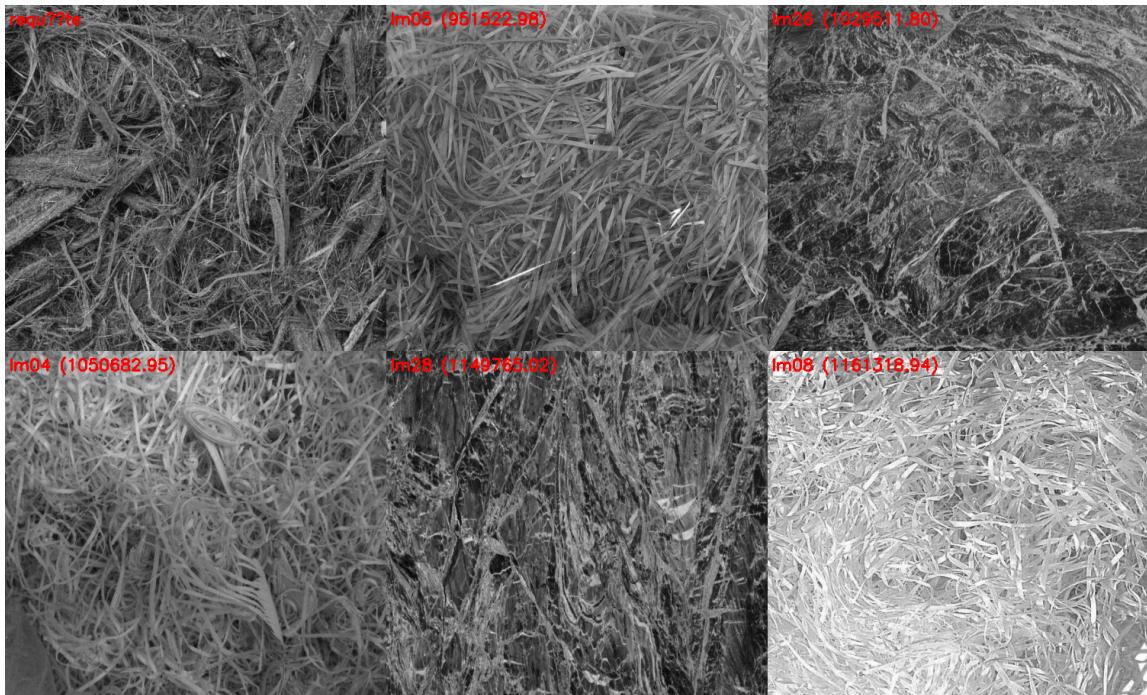


FIGURE 4 – Résultats de recherche pour l'image Im01.jpg

Le système identifie des textures présentant des caractéristiques similaires en termes de rugosité, contraste et directionnalité.

4.2.2 Requête : Im02.jpg

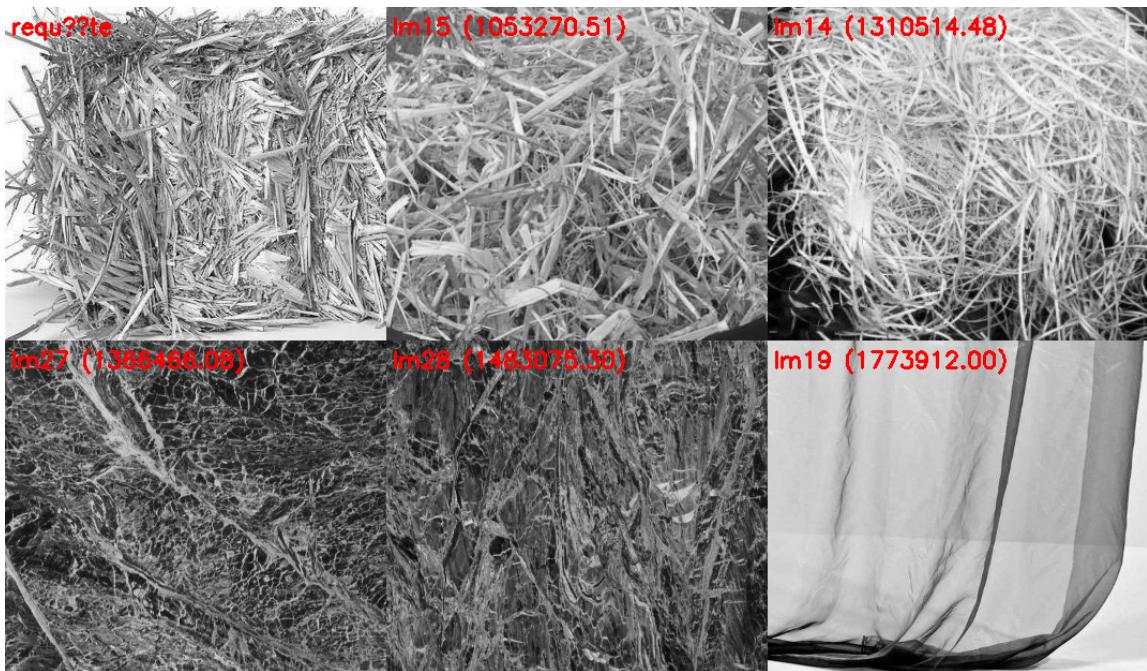


FIGURE 5 – Résultats de recherche pour l'image Im02.jpg

Les résultats montrent des textures avec des réponses similaires aux filtres de Gabor et des propriétés de Tamura comparables.

4.3 Analyse des Performances

4.3.1 Descripteurs de Forme

Les descripteurs de Fourier combinés à l'histogramme des directions offrent une représentation robuste des formes :

- **Avantages** : Invariance à la translation, rotation et échelle ; compacité de représentation (68 dimensions) ; bonne discrimination entre formes différentes
- **Limitations** : Sensibilité au bruit sur le contour ; nécessite une bonne segmentation préalable

4.3.2 Descripteurs de Texture

La combinaison des filtres de Gabor et des caractéristiques de Tamura fournit une description riche des textures :

- **Avantages** : Capture des informations multi-échelles et multi-orientations ; correspond à la perception visuelle humaine ; robustesse aux variations d'éclairage
- **Limitations** : Dimension élevée (36 caractéristiques) ; temps de calcul important pour Gabor

5 Conclusion

Ce travail a permis de développer et d'évaluer un système complet de recherche d'images par le contenu (CBIR) basé sur des descripteurs de forme et de texture. Les résultats expérimentaux obtenus démontrent l'efficacité des approches choisies pour l'indexation et la recherche d'images similaires.

5.1 Synthèse des Résultats

Pour les descripteurs de forme, la combinaison des coefficients de Fourier (32 dimensions) et de l'histogramme des directions (36 bins) offre une représentation compacte et discriminante des contours. Cette approche permet d'identifier efficacement des formes géométriquement similaires, comme l'attestent les résultats présentés sur les images de pommes, cloches et oiseaux.

Pour les descripteurs de texture, l'utilisation conjointe des filtres de Gabor multi-échelles et multi-orientations avec les caractéristiques de Tamura capture à la fois les aspects fréquentiels et perceptuels des textures. Le système réussit à retrouver des images présentant des motifs visuels similaires en termes de rugosité, contraste, directionnalité et granularité.

La mesure de similarité par distance euclidienne s'est révélée appropriée pour quantifier la proximité entre vecteurs caractéristiques, permettant un classement pertinent des résultats de recherche.

5.2 Choix Méthodologiques

Les paramètres du système ont été définis pour optimiser le compromis entre précision et efficacité computationnelle :

-
- 32 coefficients de Fourier pour capturer la forme globale et les détails significatifs
 - 36 bins pour l'histogramme directionnel, offrant une résolution angulaire de 10°
 - 16 filtres de Gabor couvrant 4 orientations et 2 échelles, pour une analyse multi-résolution
 - 4 caractéristiques de Tamura modelisant la perception visuelle humaine

5.3 Apports et Limites

Les principaux apports de ce travail incluent :

- Une implémentation fonctionnelle d'un système CBIR complet
- Une validation empirique de l'efficacité des descripteurs classiques
- Une représentation vectorielle compacte facilitant le stockage et la recherche
- Une méthodologie reproductive et extensible

Néanmoins, certaines limitations doivent être mentionnées :

- Sensibilité au bruit et à la qualité de la segmentation pour les descripteurs de forme
- Coût computationnel des filtres de Gabor pour de grandes bases d'images
- Absence de mécanismes d'apprentissage adaptatif des poids de caractéristiques
- Pas de prise en compte de la couleur dans l'analyse

5.4 Perspectives

Ce travail ouvre plusieurs perspectives d'amélioration et d'extension :

- **Fusion multimodale** : Combiner forme, texture et couleur avec pondération optimisée
- **Apprentissage profond** : Intégrer des descripteurs issus de réseaux convolutifs pré-entraînés
- **Indexation avancée** : Implémenter des structures d'indexation (KD-Tree, hashing) pour accélérer la recherche dans de grandes bases
- **Retour de pertinence** : Permettre à l'utilisateur d'affiner les résultats par feed-back itératif
- **Extension aux vidéos** : Adapter les descripteurs pour l'indexation de séquences vidéo

5.5 Conclusion Générale

Ce projet illustre l'importance de la représentation des données visuelles dans les systèmes de recherche d'images. Les descripteurs mathématiques classiques, bien que supplantés par les approches d'apprentissage profond dans certains contextes, demeurent pertinents pour leur interprétabilité, leur légèreté computationnelle et leur efficacité sur des domaines spécifiques.

Le système développé constitue une base solide pour des applications pratiques en organisation de photothèques, recherche dans des bases médicales, classification de produits, ou détection de similitudes visuelles. Il démontre que la combinaison judicieuse de descripteurs théoriquement fondés et de mesures de similarité appropriées permet d'obtenir des résultats satisfaisants en recherche d'images par le contenu.