The Multimodal Signature Method: An Introduction

D. Koubaroulis[1,2] J. Matas[1,2] J. Kittler[1]

[1]Centre for Vision Speech and Signal Processing
University of Surrey, Guildford, GU2 7XH , UK
d.koubaroulis@ee.surrey.ac.uk

[2]CMP, CTU Prague
Karlovo nám. 13, 121 35 CZ
matas@cmp.felk.cvut.cz

Abstract

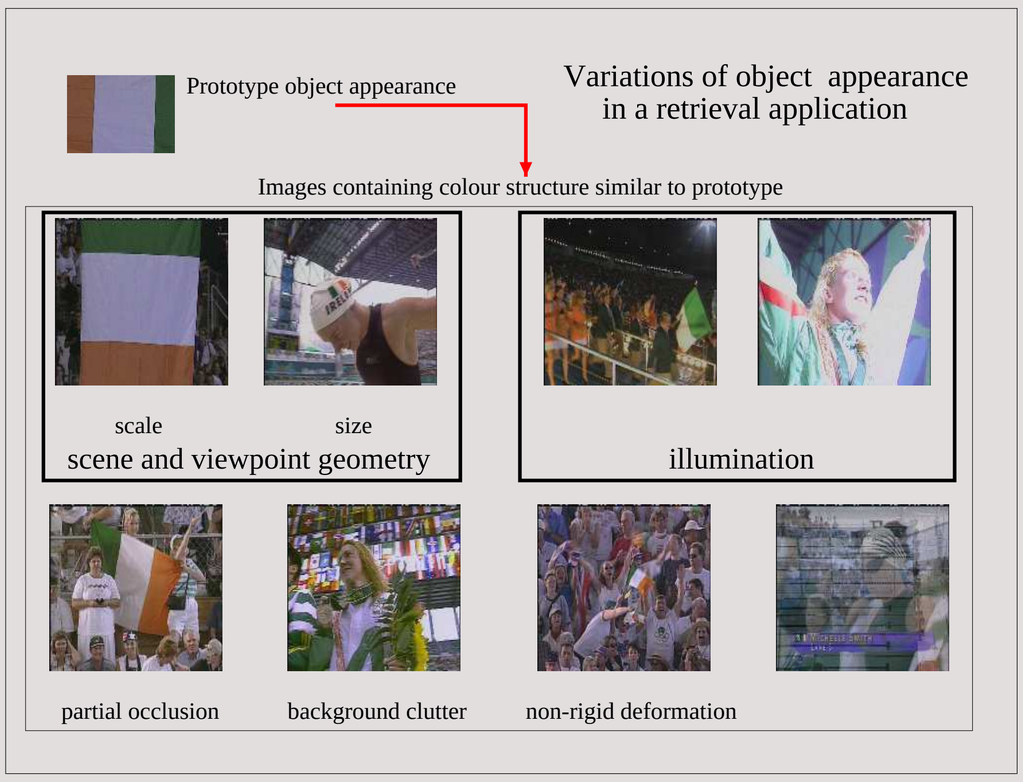
The multimodal neighbourhood signature (MNS) method has presented acceptable results both for the colour-based image retrieval and the object recognition task. Local colour content is concisely represented by invariant features computed from neighbourhoods with multimodal colour density function.

In this paper, efficiency related issues regarding the MNS algorithm are investigated. Its performance, speed, sensitivity and storage requirements are tested on a standard colour object recognition experiment. Very good recognition rate (99.9%) was achieved in real time (0.28 seconds per match). MNS signatures are concise and thus significant data reduction is achieved. The complexity of signature computation and matching are analysed and efficient implementations are proposed.

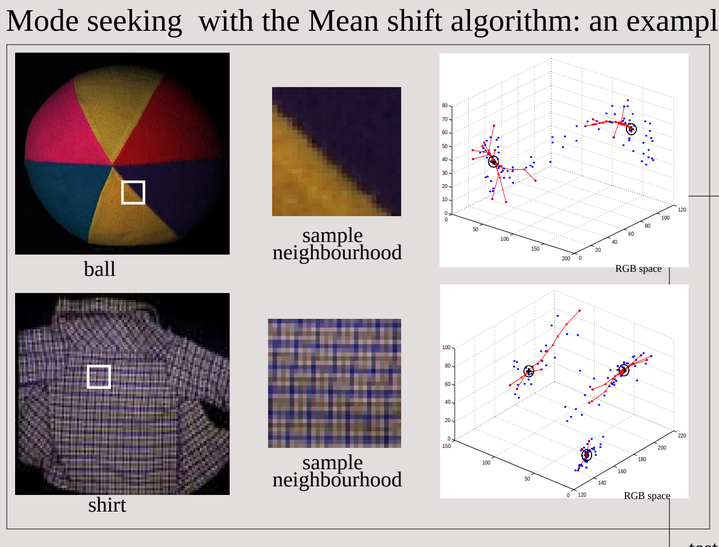
Object colour appearance (or image colour structure) is represented by stable measurements computed from local patches (neighbourhoods) with multimodal colour density function.

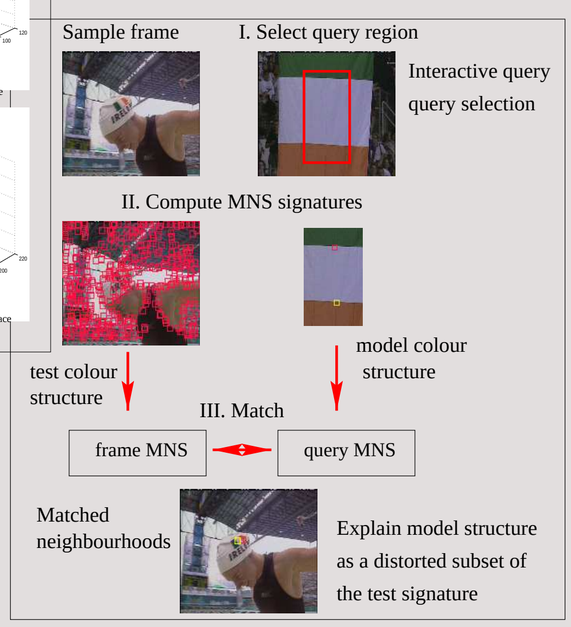
Motivation:

- colour constant descriptors cannot be computed from unimodal neighbourhoods ("Why multi-modal?")
- modes of colour density are robustly filtered, stable measurements ("Why modes?")
- local processing guarantees graceful degradation of performance w.r.t. occlusion and clutter ("Why local?")
- no need for prior segmentation or edge detection.
- local invariants computed flexibly depending on the adopted model of surface reflectance and/or illumination change. (e.g. colour measurement, chromaticities, ratios, cross-ratios)



Mode seeking with the Mean shift algorithm: an example

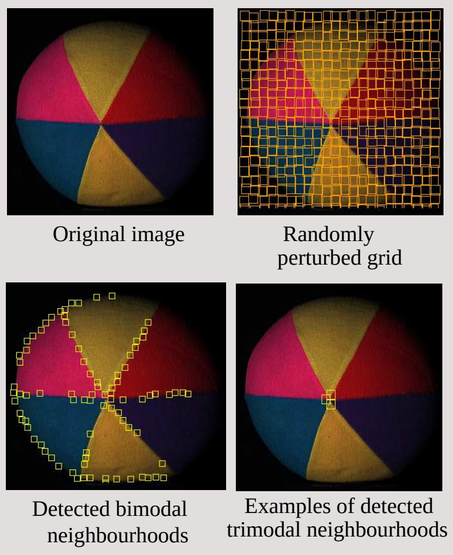




MNS-based recognition: an overview

Computing the MNS signature

1. Process neighbourhoods placed on a perturbed grid over the image.
2. For each neighbourhood locate the modes of the colour density function using the **mean shift algorithm.** [Fukunaga 75, Comaniciu and Meer 99]
3. Construct 6-dimensional RGB^2 vectors from pairs of modes of multimodal neighbourhoods.
4. Compute the MNS signature from the modes of the density function in the RGB space located with mean shift algorithm.



Algorithm 1: **Mode Seeking using the Mean Shift**

1. Select a starting data point $\bar{x}_o$.
2. Compute the mean difference vector

$$M_h(\bar{x}_o) = \frac{1}{n_{\bar{x}}} \sum_{\bar{x}_i \in S_h(\bar{x})} (\bar{x}_i - \bar{x}_o)$$

where h is the kernel radius and $n_{\bar{x}}$ is the number of data points inside the d-dimensional hypersphere S of radius h centred at $\bar{x}_o$

3. Iteratively shift the kernel by the mean shift vector ($\bar{x}'_o = M_h(\bar{x}_o)$) till convergence

- $M_h(\bar{x})$ is shown to be equal to the normalised gradient of $f(\bar{x})$ (using the Epanechnikov kernel)

$$\frac{h^2}{d+2} \frac{\hat{\nabla} f(\bar{x})}{\hat{f}(\bar{x})}$$

- **Convergence** to the closest mode is **guaranteed**. [Comaniciu99]

Objective: Establish the local maxima of the colour density function in the RGB domain

- The general kernel-based estimate of a true multivariate density function $f(\bar{x})$ at a point $\bar{x}_0$ in a d-dimensional feature space is

$$\hat{f}(\bar{x}_0) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{\bar{x}_i - \bar{x}_0}{h}\right) \quad (1)$$

where $\bar{x}_i,\ i = 1..n$ are the sample data points and K is the kernel function with kernel width h.

- Modes i.e. points $\bar{x}$ where $\nabla f(\bar{x}) = 0$ are sought. Computation of $f(\bar{x})$ is not required.
- **Mean Shift** : a simple and effective gradient ascent technique when the kernel function in (1) is the Epanechnikov kernel [Fukunaga75]

$$K_E(\bar{x}) = \begin{cases} \frac{1}{2} c_d^{-1}(d+2)(1 - \bar{x}^T \bar{x}) & \text{if } \bar{x}^T \bar{x} < 1 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where c_d is the volume of the unit d-dimensional sphere and $\bar{x}$ are the data points.

  - robust to outliers
  - minimum integrated square error

Matching MNS signatures

- Is an **assignement problem**, i.e. a problem of uniquely associating each model (query) feature to a test (database) feature.
- Implemented a modified version of the **stable marriage matching** algorithm [Gusfield89]
- A match association function $u(i) : Q \to 0 \bigcup D$ maps each model feature i to the test feature it matched or to 0. A test association function $v(j) : D \to 0 \bigcup Q$, $j \in J$.
- T_h = maximum allowed distance between two matched features.

Algorithm 1: **MNS Matching**

1. Set $u(i) = 0$ and $v(j) = 0$ $\forall i, j$.
2. From each signature s compute the invariant features f_i^M, f_j^T according to the colour change model dictated by the application.
3. Compute all pairwise distances $d_{ij} = d(f_i^M, f_j^T)$ between the model and test features.
4. Set $u(i) = j$, $v(j) = i$ if $d_{ij} < d_{kl}$ and $d_{ij} < T_h$ $\forall k, l$ with $u(k) = 0$ and $v(l) = 0$.
5. Compute signature dissimilarity as $D(s^M, s^T) = \sum_{(\forall i : u(i) \neq 0)} d_{ij} + \sum_{(\forall i : u(i) = 0)} T_h$