A Statistical Approach to Local Evaluation of a Single Texture Image

Jiří Grim, Petr Somol, Michal Haindl, Pavel Pudil

Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic, Prague

Department of Pattern Recognition
http://www.utia.cas.cz/RO

Conference PRASA, Cape Town 2005
1 Mixture Based Local Texture Model
   • Local Statistical Texture Model
   • EM Algorithm for Normal Mixtures
   • Computational Aspects of Model Estimation

2 Texture Synthesis by Local Prediction
   • Sequential Prediction by Conditional Expectation Formula
   • Examples of texture synthesis

3 Local Evaluation of the Source Texture Image
   • Local Evaluation of the Log-Likelihood Values
   • Examples of Texture Evaluation
   • Examples of Application-Oriented Texture Evaluation
   • Theoretical Aspects of the Source Texture Evaluation

4 Concluding Remarks
Digitized grey-scale texture: \( Y = [y_{ij}]_{i=1}^{I}_{j=1}, \quad y_{ij} \approx \text{grey-levels} \)

Assumption:
local statistical properties of pixels in a suitably chosen observation window are shift-invariant.

window interior: \( x = (x_1, x_2, \ldots, x_N) \in \mathcal{X}, \quad \mathcal{X} = \mathbb{R}^N \)

Method:
approximation of the joint multivariate probability density \( P(x) \) by normal mixture of product components:

\[
P(x) = \sum_{m \in \mathcal{M}} w_m F(x|\mu_m, \sigma_m) = \sum_{m \in \mathcal{M}} w_m \prod_{n \in \mathcal{N}} f_n(x_n|\mu_{mn}, \sigma_{mn})
\]

\[
f_n(x_n|\mu_{mn}, \sigma_{mn}) = \frac{1}{\sqrt{2\pi \sigma_{mn}}} \exp\left\{ -\frac{(x_n - \mu_{mn})^2}{2\sigma_{mn}^2} \right\}
\]

\( \mathcal{M} = \{1, \ldots, M\}, \quad \mathcal{N} = \{1, \ldots, N\} \approx \text{index sets} \)
EM Algorithm for Normal Mixtures

**Dataset:** $S = \{x^{(1)}, \ldots, x^{(K)}\} \approx$ by shifting observation window

$$F(x|\mu_m, \sigma_m) = \prod_{n \in N} \frac{1}{\sqrt{2\pi}\sigma_{mn}} \exp \left\{ -\frac{(x_n - \mu_{mn})^2}{2\sigma_{mn}^2} \right\}$$

**Log-likelihood criterion:**

$$L = \frac{1}{|S|} \sum_{x \in S} \log \left[ \sum_{m \in M} w_m F(x|\mu_m, \sigma_m) \right]$$

**EM Algorithm:**

$$q(m|x) = \frac{w_m F(x|\mu_m, \sigma_m)}{\sum_{j \in M} f(j)F(x|\mu_j, \sigma_j)}$$

$$f'(m) = \frac{1}{|S|} \sum_{x \in S} q(m|x), \quad \mu'_{mn} = \frac{1}{\sum_{x \in S} q(m|x) \sum_{x \in S} x_n q(m|x)} \sum_{x \in S} x_n q(m|x)$$

$$(\sigma'_{mn})^2 = -(\mu'_{mn})^2 + \frac{1}{\sum_{x \in S} q(m|x) \sum_{x \in S} x_n^2 q(m|x)} \sum_{x \in S} x_n^2 q(m|x)$$
Computational Aspects of Model Estimation

- Dimension of the window space \( N \approx 10^2 - 10^3 \)
  (e.g. for window size 20x20 pixels \( N=400 \))

- Low-frequency details \( \Rightarrow \) large window-size
  (but \( N \approx 10^3 \) becomes computationally prohibitive)

- Training data set \( S \) obtained by shifting the observation window
  (source texture image of size 500x500 pixels: \( |S| \approx 250000 \))

- Number of components \( M \approx 10^1 - 10^2 \)

- EM algorithm: random initialization, 10 - 20 iterations

- (!) Data vectors obtained by shifting the window are overlapping and
  therefore not independent

- The data set \( S \) corresponds only to a "trajectory" in \( \lambda \)
  produced by shifting (\( \Rightarrow \) less representative, bad sampling properties)
Sequential Prediction by Conditional Expectation Formula

\[ D = \{j_1, \ldots, j_l\} \subset \mathcal{N} \approx \text{defined part of the window} \]
\[ C = \{i_1, \ldots, i_k\} = \mathcal{N} \setminus D \approx \text{predicted part of the window} \]

\[ x_D = (x_{j_1}, \ldots, x_{j_l}) \in \mathcal{X}_D, \quad F(x_D|\mu_m, \sigma_m) = \prod_{n \in D} f_n(x_n|\mu_{mn}, \sigma_{mn}) \]
\[ x_C = (x_{i_1}, \ldots, x_{i_k}) \in \mathcal{X}_C, \quad F(x_C|\mu_m, \sigma_m) = \prod_{n \in C} f_n(x_n|\mu_{mn}, \sigma_{mn}) \]

conditional distribution:

\[ P_{C|D}(x_C|x_D) = \frac{P_{CD}(x_C, x_D)}{P_D(x_D)} = \sum_{m \in \mathcal{M}} W_m(x_D) F(x_C|\mu_m, \sigma_m) \]

\[ W_m(x_D) = \frac{f(m) F(x_D|\mu_m, \sigma_m)}{\sum_{j \in \mathcal{M}} f(j) F(x_D|\mu_j, \sigma_j)} \]

PREDICTION: expectation \( \bar{x}_C \) given \( x_D \)
( alternatively: random sampling by \( W_m(x_D) \) )

\[ \bar{x}_C = E_{C|D}\{x_C|x_D\} = \int x_C P_{C|D}(x_C|x_D) dx_C = \sum_{m \in \mathcal{M}} W_m(x_D) \mu_{mC} \]
**Example 1: Synthesis for Texture “Ratan”**

**Synthesis:** by random sampling the component means $\mu_{mC}$ according to the conditional weights $W_m(x_D)$:

- source texture image: 512x512 pixels $\Rightarrow |S| = 233000$
- dimension: $N = 30 \times 30 = 900$, number of components: $|M| = 80$
- number of EM iterations: $t = 15$
Local Evaluation of the Log-Likelihood Values

Motivation:
successful texture synthesis ⇒ the properties of source texture image can be described locally by the mixture model $P(x)$

**LOG-LIKELIHOOD:**
\[ \log P(x) \approx \text{measure of typicality of the window interior } x \]
**Remark:** $\log P(x)$ is highly sensitive to grey-level deviation

“Background” distribution:
\[ P_0(x) = \prod_{n \in \mathcal{N}} f_n(x_n|\mu_{0n}, \sigma_{0n}), \quad \mu_{0n} = \frac{1}{|S|} \sum_{x \in S} x_n, \quad \sigma_{0n}^2 = \frac{1}{|S|} \sum_{x \in S} x_n^2 - \mu_{0n}^2. \]

**LOG-LIKELIHOOD RATIO:**
\[ \log P(x)/P_0(x) \approx \text{structural typicality of the window interior } x \]
**Remark:** $\mu_{0n}, \sigma_{0n}$ are nearly identical for all $n \in \mathcal{N}$ ⇒ $\log P(x)/P_0(x)$ is nearly invariant with respect to grey-level deviations and it is more sensitive to structural irregularities.
**Example 1: Local Evaluation of Texture Image “Cushion”**

**Principle:** likelihood values are displayed as grey-levels at central point of the shifting observation window.

**Remark:** Log-likelihood values are highly sensitive to the deviations of grey levels. So e.g. hardly visible light pixels in the texture “cushion” (left image) produce dark spots of window size (central image).
Example 2: Local Evaluation of Texture Image “Ratan”

**Principle:** likelihood values displayed as grey-levels at central point of the shifting observation window

**Remark:** Log-likelihood ratio is less dependent on the grey-level mean and it is more sensitive to structural differences. The structural irregularities of the “ratan” texture (cf. left image) are therefore more clearly identified by the log-likelihood ratio (right image) than by the log-likelihood alone (central image).
**Principle:** reversely displayed log-likelihood values (central image) and log-likelihood ratio values (right image) in the red spectral component
**Principle:** reversely displayed log-likelihood values (central image) and log-likelihood ratio values (right image) in the red spectral component
Theoretical Aspects of the Source Texture Evaluation

- unlike other fields (like texture modelling, prediction, pattern recognition) the estimated mixture \( P(x) \) is applied to the “training” data set \( S \) again
- log-likelihood criterion optimally “fits” the estimated mixture \( P(x) \) to the data set \( S \)
- \( \Rightarrow \) application of the mixture model \( P(x) \) to the source data \( x \in S \) is well justified by the estimation procedure
- \( \Rightarrow \) log-likelihood value is a suitable measure of typicality of data vectors \( x \in S \)
- limited representativness of the set \( S \) is less relevant since \( P(x) \) is not applied to the data not contained in \( S \)
Concluding Remarks

Local Evaluation of a Single Texture Image:

- identifies abnormal (unusual) texture pieces to be analyzed in detail
- it is non-supervised, need not be trained by other images
- it is useful in case of a large variability of evaluated textures (e.g. in medical diagnostics)
- the result of evaluation has a clear interpretation

Application possibilities:

- fault detection
- abnormality or novelty detection
- mammographic screening
- monitoring of complex processes (to detect exceptional states)


Example 2: Synthesis for Texture “Fabrik”

**Synthesis:** by random sampling the component means $\mu_{mC}$ according to the conditional weights $W_m(x_D)$

- **original image**
- **sampling of means**
- **“centroid” sampling**

**right image:** component means replaced by similar pieces of the original texture ("centroids", "micro-tiles") ⇒ stochastic tiling

- source texture image: 512x512 pixels ⇒ $|S| \approx 232000$
- dimension: $N = 30 \times 30 = 900$, number of components: $|M| = 90$
- number of EM iterations: $t = 20$
**Example 3: Synthesis for Texture “Leather”**

**Synthesis:** by random sampling the component means $\mu_{mc}$ according to the conditional weights $W_m(x_D)$

![Original Image](image1)
![Sampling of Means](image2)
![“Centroid” Sampling](image3)

**Right Image:** component means replaced by similar pieces of the original texture (“centroids”, “micro-tiles”) ⇒ stochastic tiling

- source texture image: 512x512 pixels ⇒ $|S| = 242000$
- dimension: $N = 20 \times 20 = 400$, number of components: $|\mathcal{M}| = 50$
- number of EM iterations: $t = 15$
**Principle:** reversely displayed log-likelihood values (central image) and log-likelihood ratio values (right image) in the red spectral component
Example 4: Irregularity Evaluation - Texture “Flowers”

**Principle:** reversely displayed log-likelihood values (central image) and log-likelihood ratio values (right image) in the red spectral component.
**Example 5: Irregularity Evaluation - Texture “Carpet”**

**Principle:** inversely displayed log-likelihood values (central image) and log-likelihood ratio values (right image) in the red spectral component.