Segmentation-Based Shape-Adaptive Image Coding

Zhao Ying

Supervised by Professor D. Malah

Department of Electrical Engineering

Technion-Israel Institute of Technology

27 April, 1999

Purpose & Motivation

- Block-based image coding systems (JPEG) produce poor quality images at low bit-rate (<0.5 bpp)
- HVS is sensitive to distortion along object boundary, less sensitive to distortion within the content
- Possible approach proposed: segmentation-based image coding, including
  - Image segmentation system
  - Texture coding system
  - Contour coding system

Outline

- Gray-Scale Morphology - A Brief Overview
- Segmentation Algorithms
- Shape-Adaptive Image Coding
  – Region-Based
  – Block-Based
- Conclusion & Future Studies

Gray Scale Morphological Dilation & Erosion

- Dilation

$$\delta_n(f)(x) = Max\{f(x-y), y \in M_n\}$$

- Erosion

$$\varepsilon_n(f)(x) = Min\{f(x+y), y \in M_n\}$$

chap21.eps

M_n - Structuring element of size n

7×7 square used in the example ($n = 3$)

Image of size 100×100

Reconstruction by Dilation (Erosion)

- Geodesic Dilation Size 1

$$\delta^1(f,r) = Min\{\delta_1(f), r\}$$

- Geodesic Erosion Size 1

$$\varepsilon^1(f,r) = -\delta^1(-f,-r)$$

- Reconstruction by Dilation

$$\gamma^{(rec)}(f,r) = \delta^\infty(f,r) = (\cdots\delta^1(\cdots\delta^1(f,r)\cdots,r))$$

- Reconstruction by Erosion

$$\phi^{(rec)}(f,r) = \varepsilon^\infty(f,r) = (\cdots\varepsilon^1(\cdots\varepsilon^1(f,r)\cdots,r))$$

Morphological Filters & Operators

- Morphological Opening $\gamma_n(f) = \delta_n(\varepsilon_n(f))$
- Morphological Closing $\phi_n(f) = \varepsilon_n(\delta_n(f))$
- Opening by Reconstruction of Erosion $\gamma^{(rec)}(\varepsilon_n(f), f)$
- Closing by Reconstruction of Dilation $\phi^{(rec)}(\delta_n(f), f)$
- h_maxima $h\_\max(f) = \gamma^{(rec)}(f - h, f)$
- h_minima $h\_\min(f) = \phi^{(rec)}(f + h, f)$
- Morphological Watershed

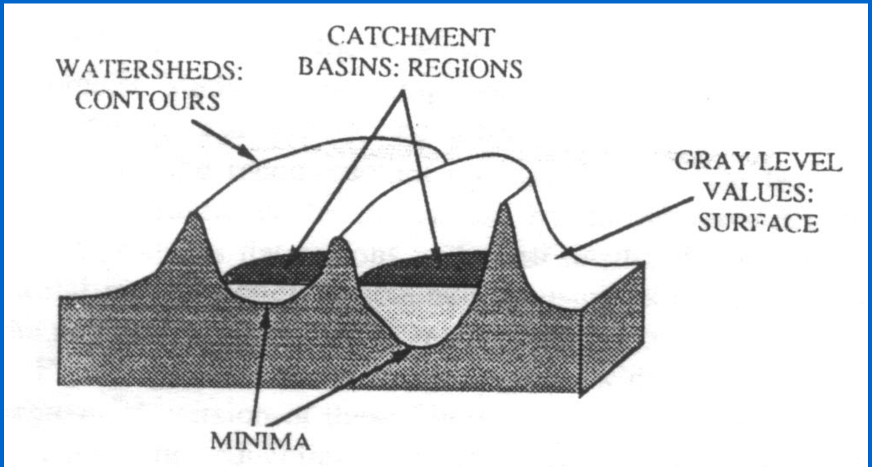
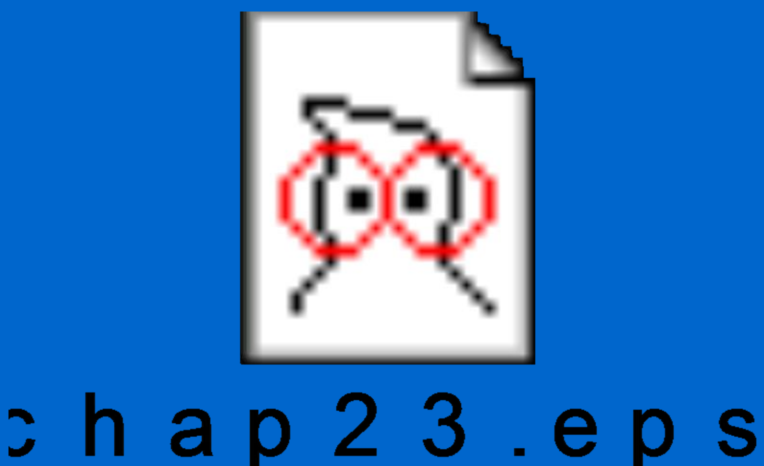


Image Simplification by Morphological Filters (Operators)

Image of size 512×512

25×25 square structuring element



Morphological-Based Image Segmentation Algorithms

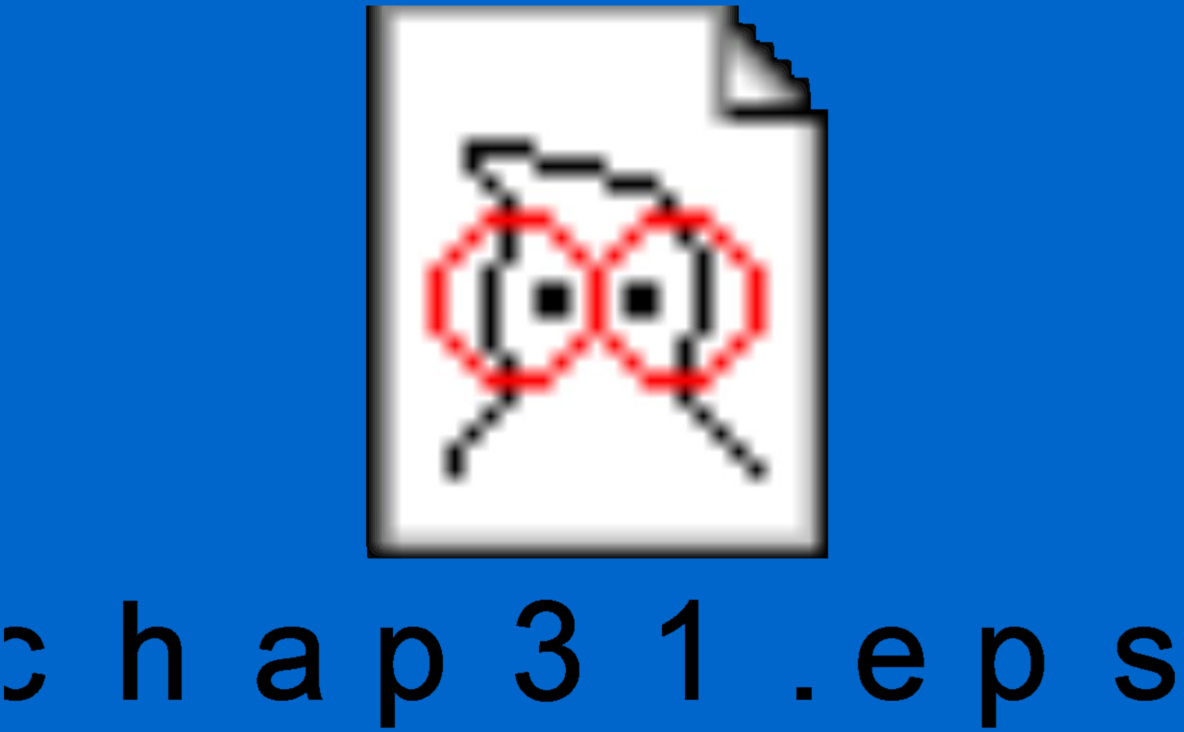
Existing

- Hierarchical Morphological Synthesis-by-Analysis
- Morphological Simplification, Region Splitting & Merging

Proposed

- Edge Detection, Local-Activity Classification

Hierarchical Morphological Synthesis-by-Analysis Segmentation Algorithm [P.Salembier, 1995]



Hierarchical Morphological Synthesis-by-Analysis Segmentation Results

chap35.eps

- Proper size of the filters is unknown
- Watershed produces false contours
- High complexity, high computation load

Morphological Simplification, Region Splitting & Merging Segmentation Algorithm [D. Wang, 1997]

chap32.eps

Morphological Simplification, Region Splitting & Merging Segmentation Results

- Oversegmented output
- Complicated contour is produced

Edge Detection, Local-Activity-Classification Segmentation Algorithm

Local-Activity Classification [J. Pandel, 1991]

Apply operators $P(k),\ k=1,\cdots,4,$
to image to calculate local-activity:

$$A_{x,y}(k)=\sum_{i=1}^{5}\sum_{j=1}^{5} f(x-3+i,y-3+j)\,p_{(i,j)}(k)$$

$$A_{x,y}=\max_{k=1,\cdots,4}\left(\left|A_{x,y}(k)\right|\right)$$

P(1)

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{matrix}$$

P(2)

$$\begin{matrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{matrix}$$

P(3)

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{matrix}$$

P(4)

$$\begin{matrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{matrix}$$

class 1: $0 \le A_{x,y} < 32$: flat areas

class 2: $32 \le A_{x,y} < 64$: low structured regions

class 3: $64 \le A_{x,y} < 128$: high structured regions

class 4: $128 \le A_{x,y}$: edges

Edge Detection, Local-Activity-Classification Segmentation Steps

chap37.eps

Contour Simplification by Majority Filter [C. Gu, 1994]

- Majority operator $M(P(x))$ assigns the majority label within the neighborhood to current pixel
- Apply majority operator iteratively, until no change is caused $\sigma_B = M^{\infty}(P(x))$
- Size of the filter is optional

chap38.eps

Edge Detection, Local-Activity-Classification Segmentation Results

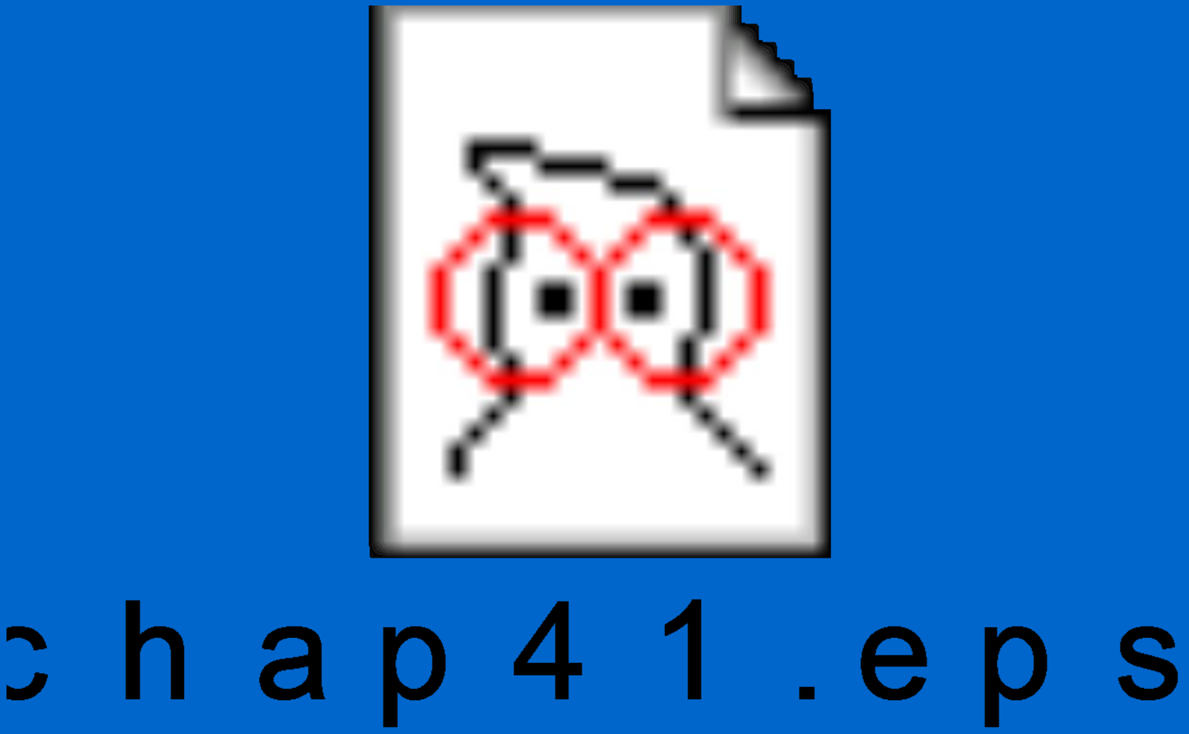
chap39.eps

chap310.eps

Advantages of Proposed Segmentation Algorithm

- Segmentation output better matched to coding task
- Moderate computation load
- Possible approach to generic segmentation algorithm

Shape-Adaptive Image Coding System



Region-Based Image Coding Techniques

- Least Squares Approximation
- Shape-Adaptive Orthogonal Basis Function Generation
  - Gram-Schmidt orthogonalization scheme
  - Householder polynomials [M. Weisfeld, 1959]
  - Weakly separable basis functions [W. Philips, 1996]

Drawbacks of Orthogonal Basis Functions Approximation

- Expensive computation
- Improper generation caused by initial dependent functions
- Unpredictable coefficients positions in frequency domain
- Limited degree can be generated

Results of Basis Function Approximation

chap52.eps

chap53.eps

Block-Based Shape-Adaptive Approach

Block Classification

- Inner blocks
- Boundary blocks
- Outside blocks

chap45.eps

Block-Based Shape-Adaptive Coding Techniques

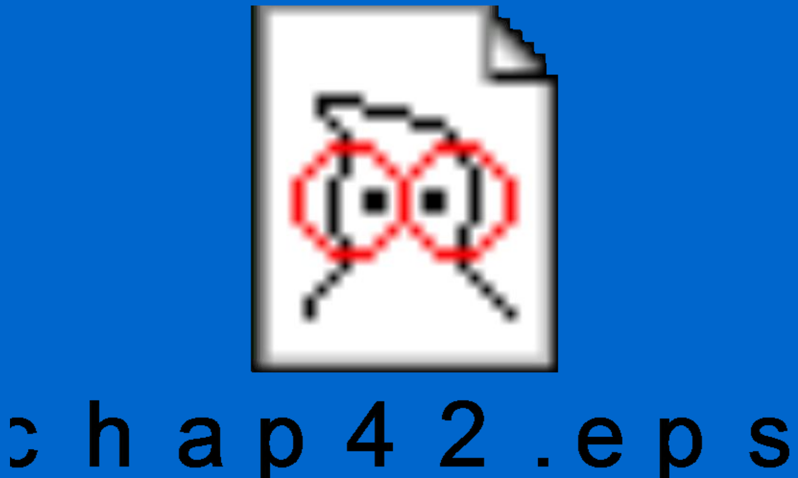
- Shape-Adaptive DCT (SADCT)
- Block Extrapolation Methods
  - Low-pass extrapolation (LPE) padding technique
  - Optimal data extrapolation (min. norm sense)

$$\text{min. } \|l\|_2 \text{ (Method of Frames)}$$

$$\text{min. } \|l\|_1 \text{ (Basis Pursuit)}$$

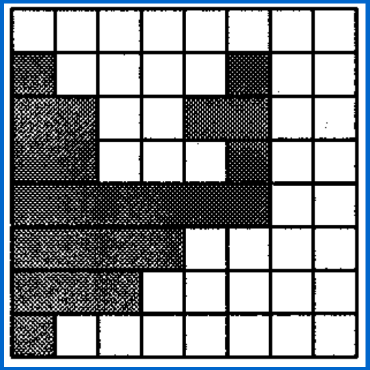
Shape-Adaptive DCT

- Shift and align data column wise (Y’), calculate 1D-DCT adaptive to column size (U)
- Shift and align DCT coefficients row wise (X’), calculate 1D-DCT adaptive to row size (V)
- Conduct reverse procedure SA-IDCT for reconstruction



Shape-Adaptive DCT (cont'd)

- Comparatively low computation load
- Less energy compaction than with 2D-DCT
- Inefficient coding of blocks having non-contiguous data



Low-Pass Extrapolation Padding

- **Algorithm Description**
  - Calculate mean of block-pixels within the segment
  - Assign the mean to block-pixels outside the segment
  - Filter each outside block-pixel by the operator

$$\tilde{f}(i,j) = [f(i,j-1) + f(i-1,j) + f(i,j+1) + f(i+1,j)]/4$$

  - Apply filtering step iteratively, until no data change
- **Properties**
  - Low computation requirements
  - Insufficient reduction of bit-rate

Optimal Block Extrapolation (min. norm)

- From 2D to 1D

$$C^T A C = S \Leftrightarrow \Phi \underline{a} = \underline{s} = \begin{bmatrix} \underline{\beta}_{m\times 1} \\ \underline{\theta}_{(N-m)\times 1} \end{bmatrix}; \ \Phi = C^T \otimes C$$

- Optimal Approach

$$\min \|\underline{a}\|_p \quad \text{subject to} \quad \Phi'_{m\times N}\, \underline{a} = \underline{\beta}$$

  - $\min \|\underline{a}\|_2$ subject to $\Phi' \underline{a} = \underline{\beta}$ (Method of Frames)
  - $\min \|\underline{a}\|_1$ subject to $\Phi' \underline{a} = \underline{\beta}$ (Basis Pursuit)

l_1 Optimal Block Extrapolation

$$\underline{a} = \underline{u} - \underline{v} \quad \underline{u},\ \underline{v} \geq 0$$

$$\min 1^T \begin{bmatrix} \underline{u} \\ \underline{v} \end{bmatrix} \text{ subject to } \begin{bmatrix} \Phi' & -\Phi' \end{bmatrix} \begin{bmatrix} \underline{u} \\ \underline{v} \end{bmatrix} = \underline{\beta}$$

- Can be solved by Linear Programming
- Only *m* nonzero entries exist in $\underline{a}$, the same number as of known original data
- Positions of nonzero entries are unpredictable

Block Data Extrapolation

chap43.ps

Block Data Transform

Use $p_i = \dfrac{c_i^2}{\sum_k c_k^2}$, $E = -\sum_i p_i \log p_i$, $i = 0, \cdots, N-1$

to calculate entropy E:

MOF: $E = 2.6489$

SADCT: $E = 0.5630$

LPE: $E = 0.2181$

BP: $E = 0.5129$

Block-Based Shape-Adaptive Coding Results

chap47.eps

chap48.eps

chap4t1.eps

Block-Based Shape-Adaptive Coding Results (cont'd)

chap4t2.eps

chap4t3.eps

Conclusions & Future Studies

- **Segmentation**
  - * Proposed algorithm is better matched to the coding task
  - * Proposed algorithm could be the basis for a generic segmentation algorithm, provided that region refinement is further studied
- **Shape-Adaptive Coding**
  - – Region-based approach has practical (numerical) problems concerning the order of basis functions can be generated

Conclusions & Future Studies (cont'd)

- **Shape-Adaptive Coding**

  * Block-based approach reduces bit-rate (compared to JPEG) for images with steep edges between segments

  * Compatible with available block-based coding system

  – Does not fully exploit the potential provided by segmentation

  BP Extrapolation

  * Provides better results than LPE padding & SADCT

  – High complexity

Least Squares Approximation

$$g(x,y) = a_1\phi_1(x,y) + a_2\phi_2(x,y) + \cdots + a_N\phi_N(x,y)$$

$$d(f,g) = \sum_{(x,y)\in D_M} (g(x,y) - f(x,y))^2$$

$$\sum_{n=1}^{N} a_n \sum_{(x,y)\in D_M} \phi_n(x,y)\phi_q(x,y) = \sum_{(x,y)\in D_M} f(x,y)\phi_q(x,y), \text{ for } q = 1,\cdots,N$$

Orthogonal Basis Functions with respect to Segment Shape

$$\sum_{(x,y)\in D_M} \phi_n(x,y)\phi_q(x,y) = 0, \text{ for } n \neq q$$

Least Squares Approximation (cont'd)

$$a_q = \frac{\sum_{(x,y)\in D_M} f(x,y)\phi_q(x,y)}{\sum_{(x,y)\in D_M} \phi_q(x,y)\phi_q(x,y)}, \text{ for } q = 1,\cdots,N$$

- Orthogonal functions are guaranteed by linear independence of initial functions
- Coefficients' calculation is numerically stable and simple
- No need to recalculate previous coefficients when new coefficients are need

Block-Based Shape-Adaptive Coding Conclusions

- BP extrapolation reduces bit-rate for images with big smooth regions, and steep edge exists within the segments
- The compression advantage of BP extrapolation coding over LPE padding and SADCT is obvious
- Block-based approach is compatible with available coding systems
- BP extrapolation coding needs non-standard VLC tables
- The complexity of BP extrapolation is high
- The potential for compression provided by segmentation is not fully exploited

JPEG Baseline Coding System