

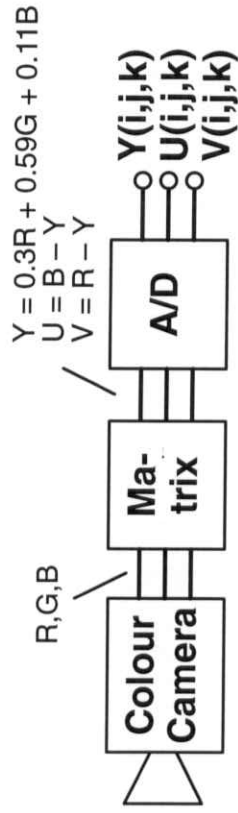
# OBJECT-BASED ANALYSIS-SYNTHESIS CODING

Jörn Ostermann  
AT&T Bell Laboratories  
Holmdel, NJ, USA

1. Source Models in Image Coding
2. Block-Based Coding
3. Object-Based Analysis-Synthesis Coding (OBASC)
4. Knowledge-Based Coding
5. Conclusion

Research was mainly performed at the Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung Universität Hannover

# Digital Television Format acc. to CCIR Rec. # 601



Luminance Y                      Chrominance U, V

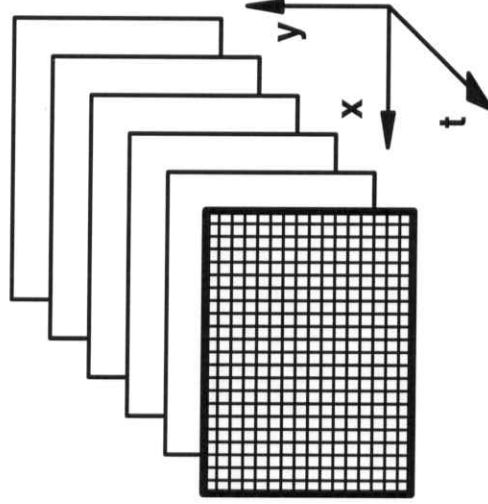
50	fields/s	50
288	lines/field	288
720	pixels/line	360
8	bit/pixel	8

82.9 MBit/s                      2x41.5

166 MBit/s

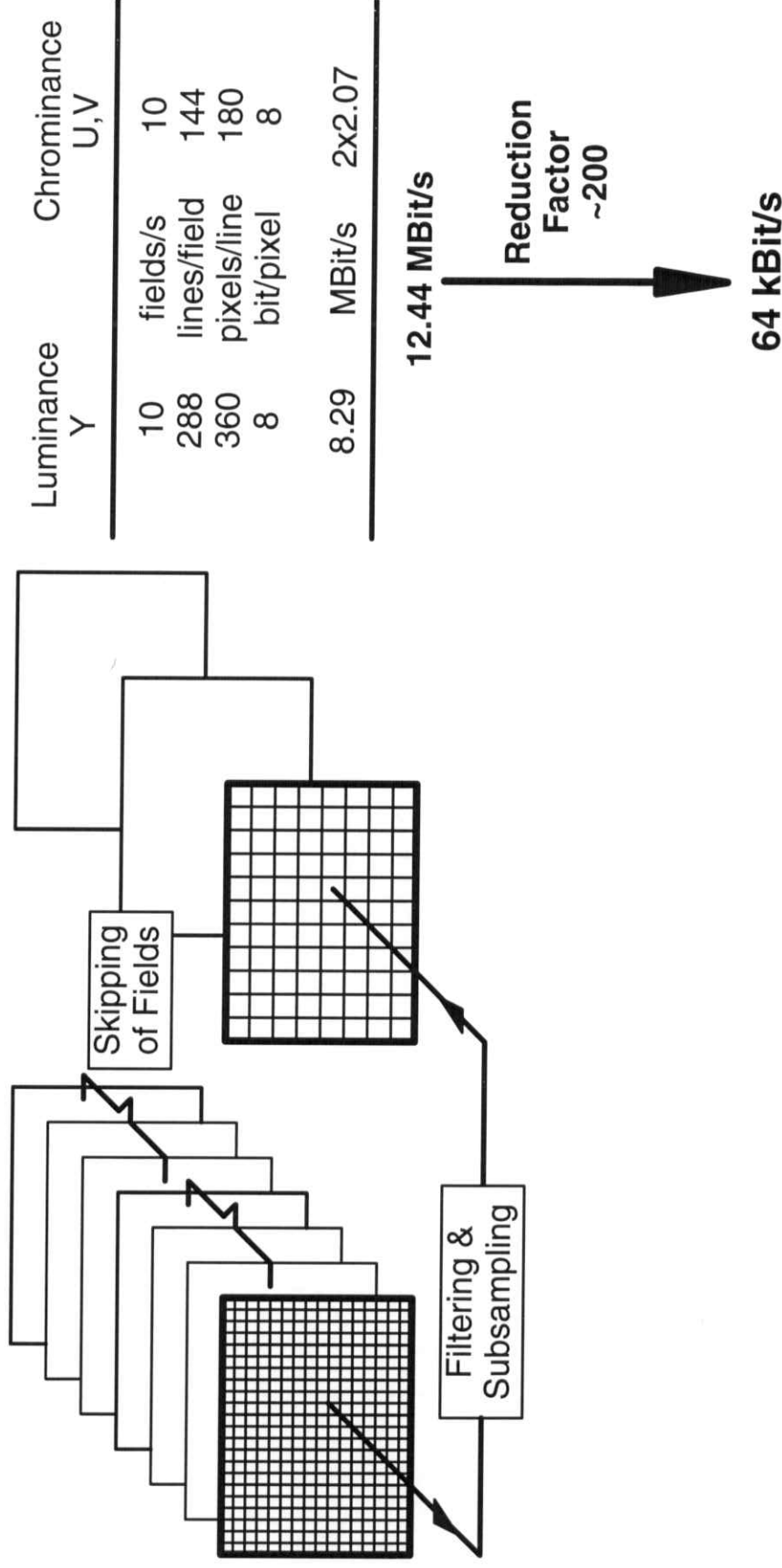
Reduction Factor  
~2600

64 kBit/s



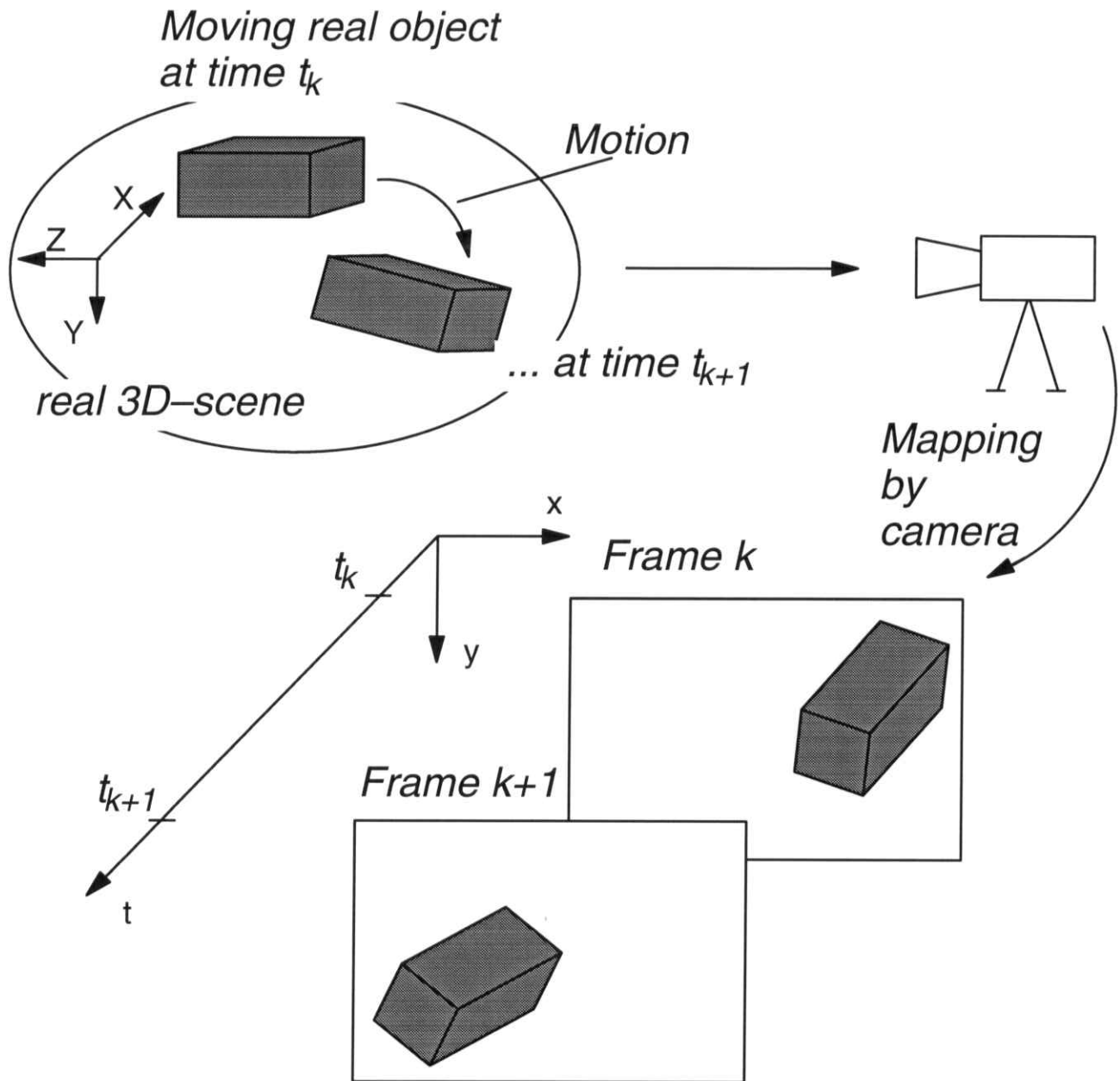
MPEG 2: 4 MBit/s

# Reduced Video Format used for Video Telephony (CIF)



MPEG 1: 1 MBit/s

# Generation of Video



Source Model:

Description of the Frame–Differences  
by Motion of Model Objects

# Use and Efficiency of Source Models

---

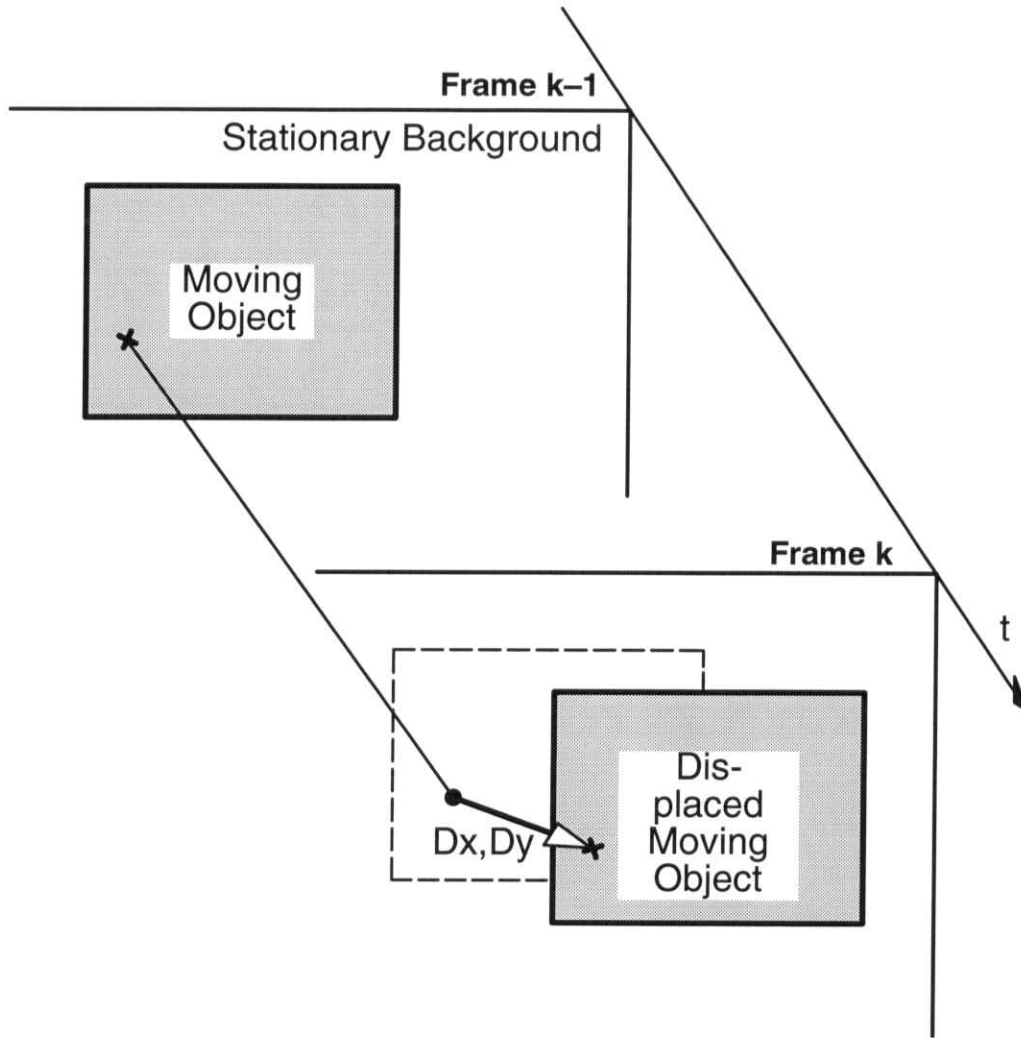
1. Use of a source model
  - to describe the image signal by parameters of the model
  - to encode the model parameters instead of the original image signal
  - redundancy reduction
  - irrelevance reduction
  
2. Efficiency of a source model
  - can be measured with the data rate required for encoding the model parameters

SOURCE MODEL	ENCODED INFORMATION	CODING TECHNIQUE
PELS	COLOUR OF PELS	PCM
STATISTICALLY DEPENDENT PELS	COLOUR OF PELS OR BLOCK OF PELS	PREDICTIVE CODING TRANSFORM CODING
TRANSLATORICALLY MOVING BLOCKS	COLOUR OF BLOCKS AND MOTION VECTORS	MOTION COMPENSATED HYBRID CODING
MOVING UNKNOWN OBJECTS	SHAPE, MOTION AND COLOUR OF EACH OBJECT	OBJECT-BASED ANALYSIS-SYNTHESIS CODING
MOVING KNOWN OBJECT	SHAPE, MOTION AND COLOUR OF THE KNOWN OBJECT	KNOWLEDGE-BASED CODING
FACIAL EXPRESSIONS	ACTION UNITS	SEMANTIC CODING

SOURCE MODEL	ENCODED INFORMATION	CODING TECHNIQUE
PELS	COLOUR OF PELS	PCM
STATISTICALLY DEPENDENT PELS	COLOUR OF PELS OR BLOCK OF PELS	PREDICTIVE CODING TRANSFORM CODING
TRANSLATORICALLY MOVING BLOCKS	COLOUR OF BLOCKS AND MOTION VECTORS	MOTION COMPENSATED HYBRID CODING
MOVING UNKNOWN OBJECTS	SHAPE, MOTION AND COLOUR OF EACH OBJECT	OBJECT-BASED ANALYSIS-SYNTHESIS CODING
MOVING KNOWN OBJECT	SHAPE, MOTION AND COLOUR OF THE KNOWN OBJECT	KNOWLEDGE-BASED CODING
FACIAL EXPRESSIONS	ACTION UNITS	SEMANTIC CODING

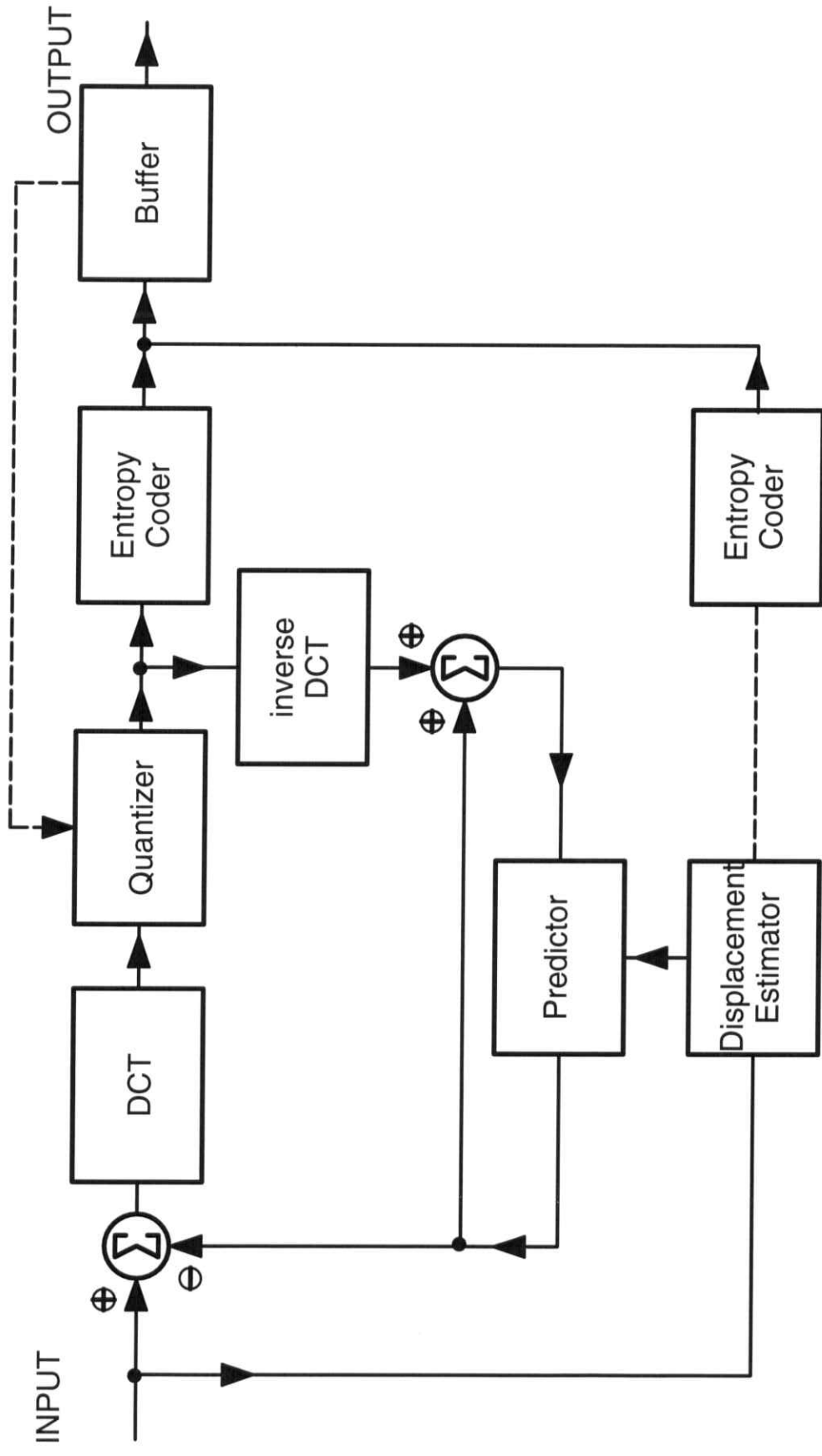
# Object Displacement and Motion Compensated Prediction

---





# Block Diagram of a Hybrid Image Sequence Coder



H.261, MPEG-1, MPEG-2, H.262, H.263

# Source Models in Image Coding

---

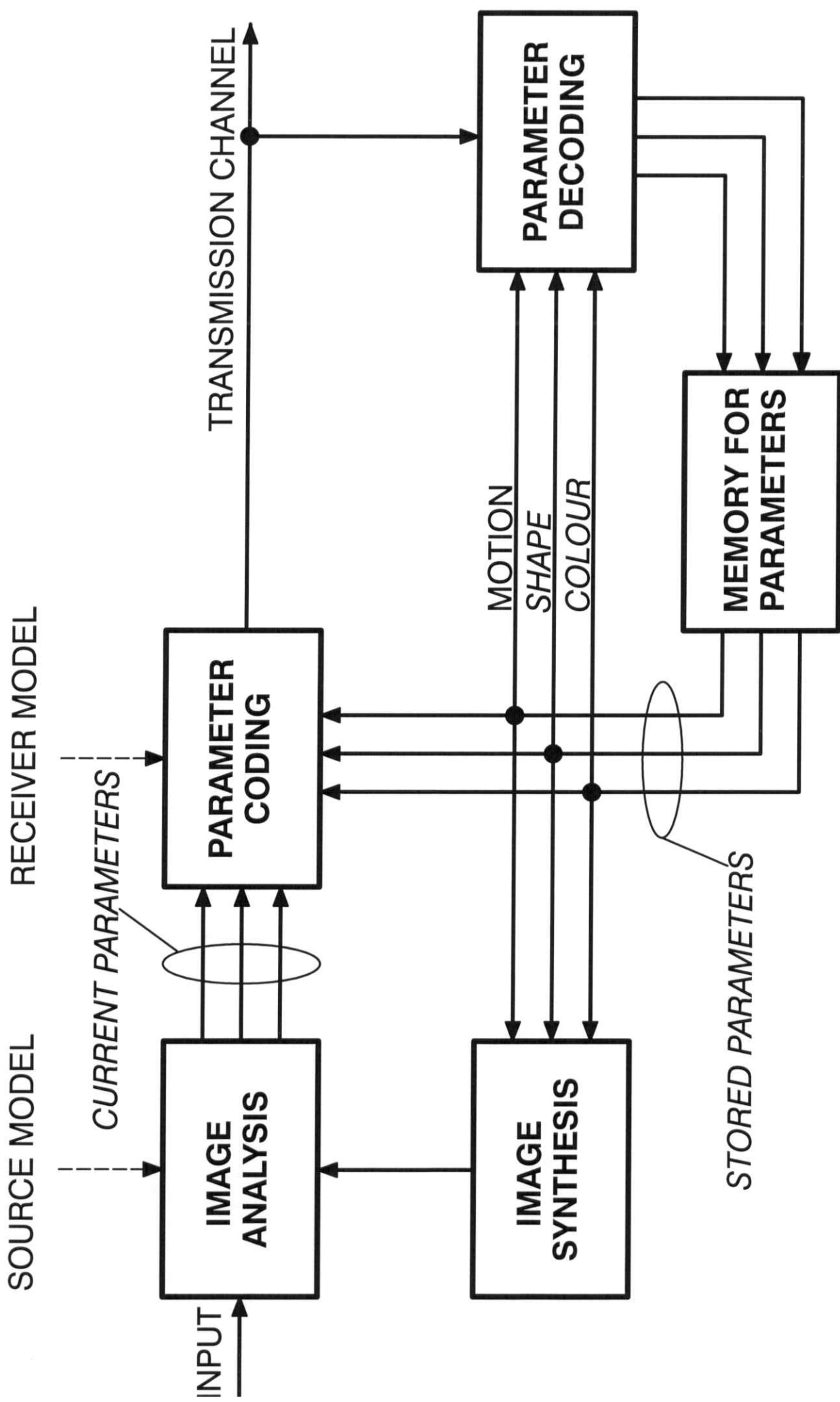
## H.261/MPEG1/MPEG2:

- Source model: Translatorically moving square blocks
- Model parameters: Motion, colour
- Limitations:
  - ◆ boundaries of moving objects
  - ◆ prediction image subjectively unpleasant
  - ◆ low irrelevance reduction

## Object–Based Analysis–Synthesis Coding (OBASC):

- Source models: Moving 2D flexible objects / moving 3D objects
- Model parameters: Motion, colour and SHAPE
- Advantages:
  - ◆ boundaries of moving objects
  - ◆ subjectively correct prediction image
  - ◆ higher irrelevance reduction
- Disadvantage:
  - ◆ additional parameter set

# Block Diagram of an Object-Based Analysis-Synthesis Coder



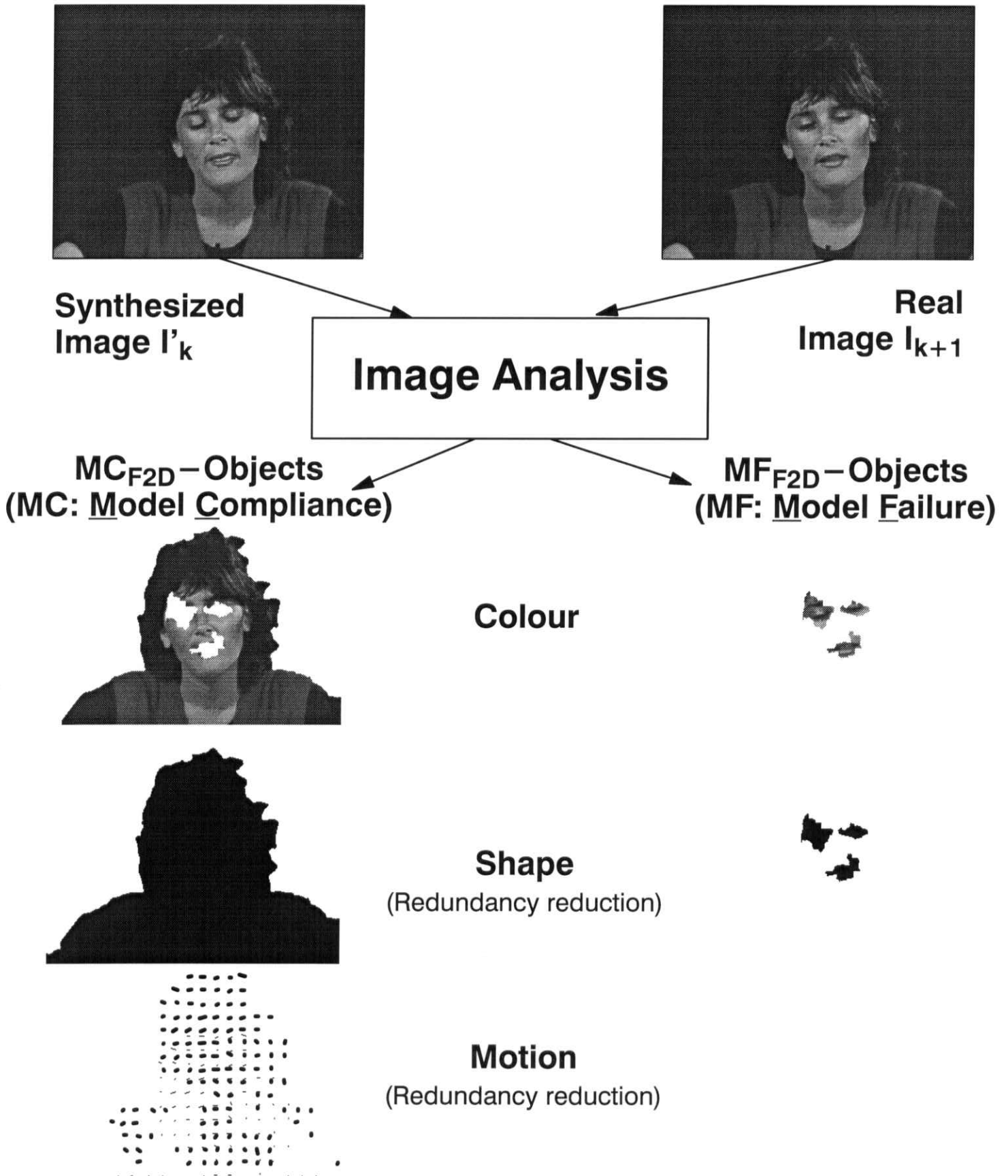
# Source Models for OBASC

---

- 2D–Objects with 2D Motion
  - Rigid 2D–Objects
  - Flexible 2D–Objects
- 2D–Objects with 3D–Motion
- 3D–Objects with 3D Motion
  - Rigid 3D–Objects
  - Flexible 3D–Objects

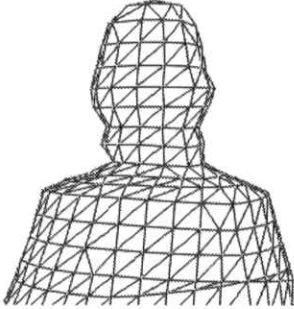
# Image Analysis

## Source Model: Flexible 2D–Object (F2D)



# Source Models for 3D-Objects: Rigid Object (R3D) I

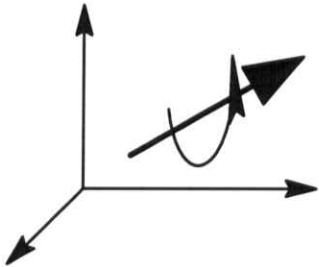
---



Shape  $\vec{F}$ : N control points  $P$



Texture  $\vec{C}$ :  $s_k$



Motion  $\vec{B}$ :  $R_x, R_y, R_z, T_x, T_y, T_z$

$$P' = [R] \cdot P + T$$

# Source Models for 3D–Objects: Rigid Object (R3D) II

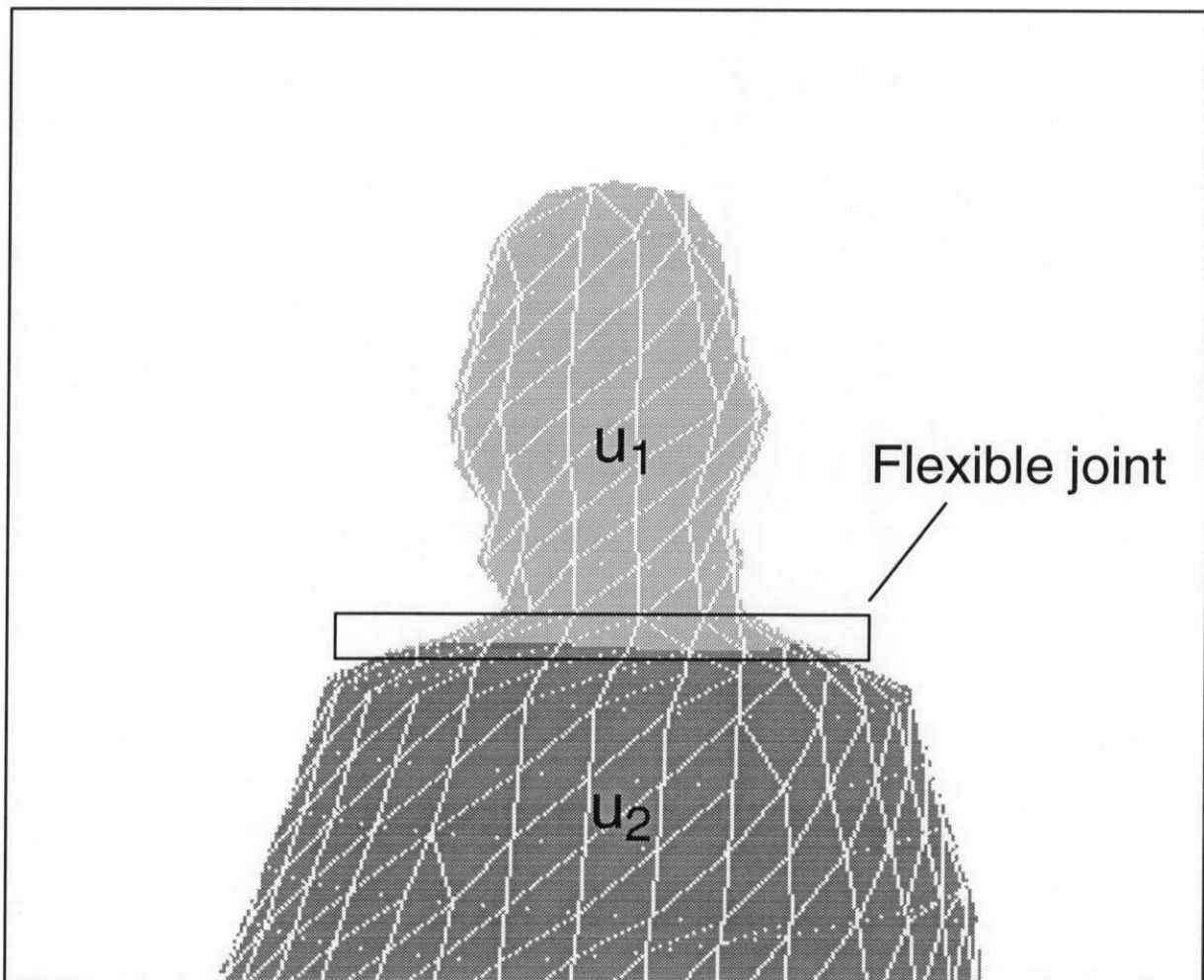
---

Shape of real objects:

- Flexibly connected rigid components

Shape of model objects:

- Flexibly connected rigid model object components  $u$ , motion parameters  $\vec{B}_u$

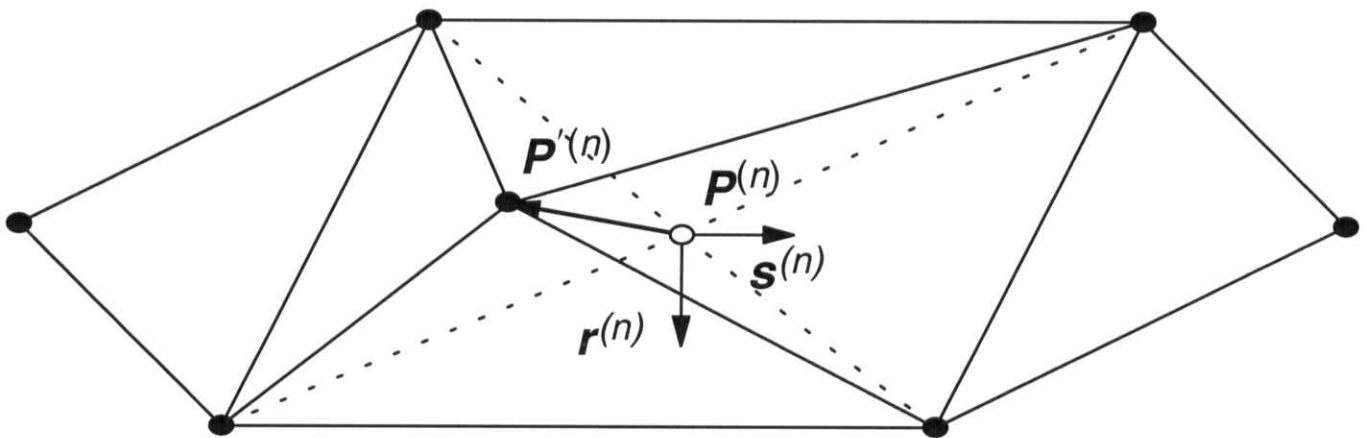


# Source Models for 3D-Objects: Flexible Objekte (F3D)

---

Extension of source model R3D

Flexible shift of control points in plane on the surface



2 shift parameters/control point

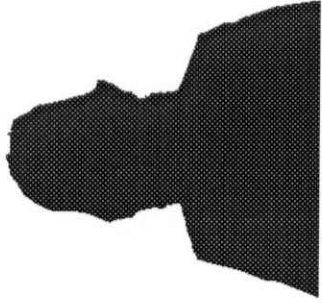
$$\mathbf{P}'^{(n)} = \mathbf{P}^{(n)} + s_r^{(n)} \mathbf{r}^{(n)} + s_s^{(n)} \mathbf{s}^{(n)}$$



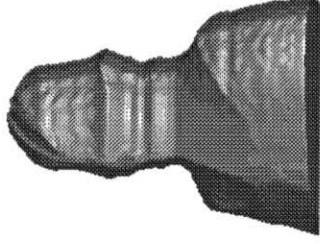
# Modelling 3D – Shape from an Object Silhouette

---

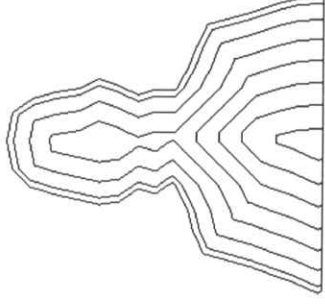
Silhouette



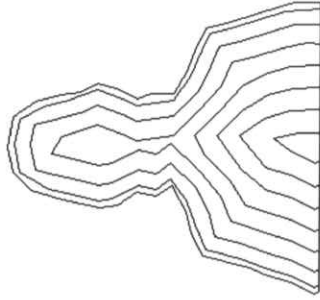
3D – Shape



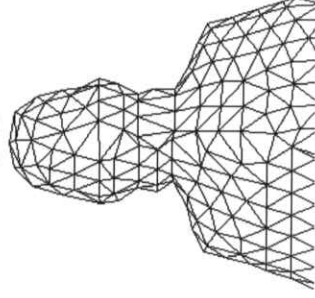
Contour Lines



Polygons



Interconnection of Polygons

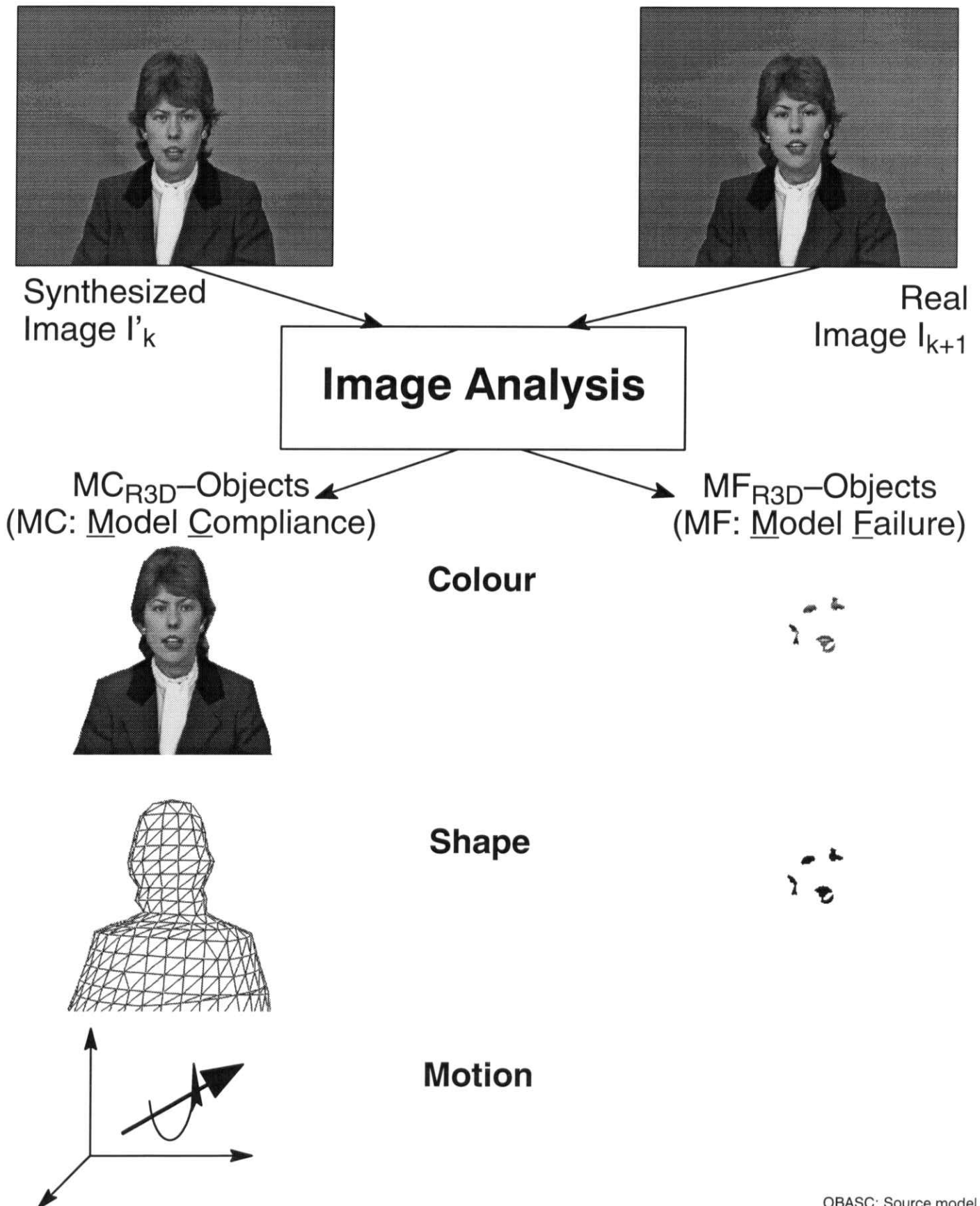


Model Object



# Image Analysis

## Source Model: Rigid 3D-Object (R3D)

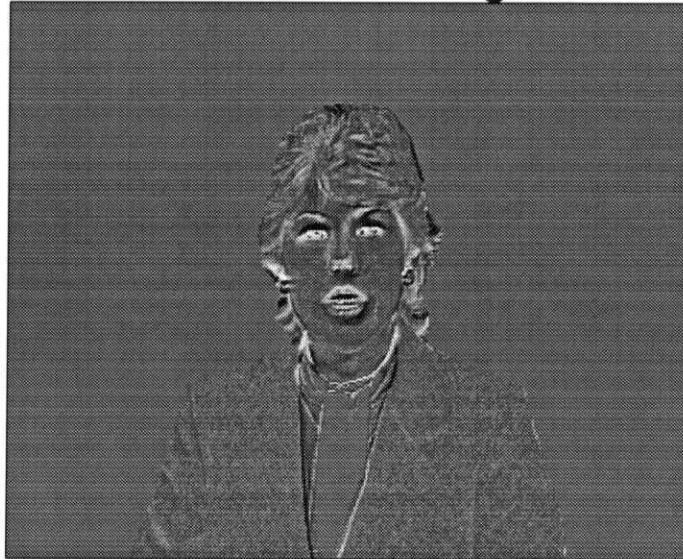


# Detection of Areas of Model Failure I

---

**Synthesis Error:** areas where real image and synthesized image differ significantly when compared to background noise

Difference Image



Areas of Synthesis Errors



# Detection of Areas of Model Failure II

---

Areas of synthesis errors

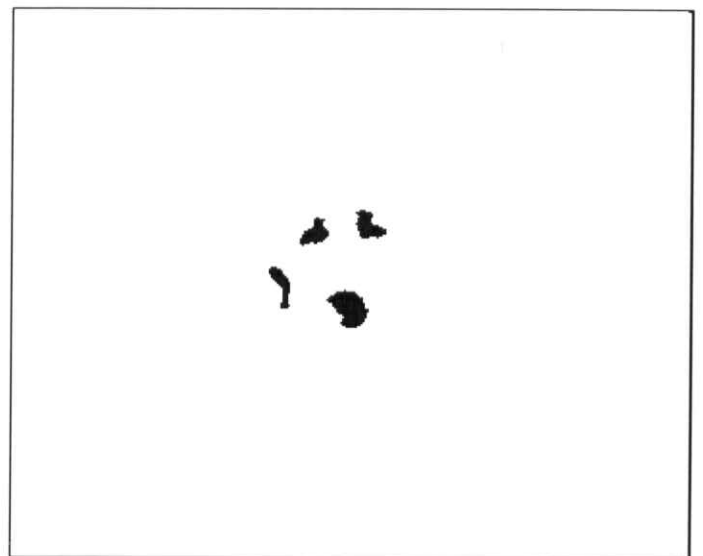


error signal  
size of area  
shape of area

Geometrical distortions

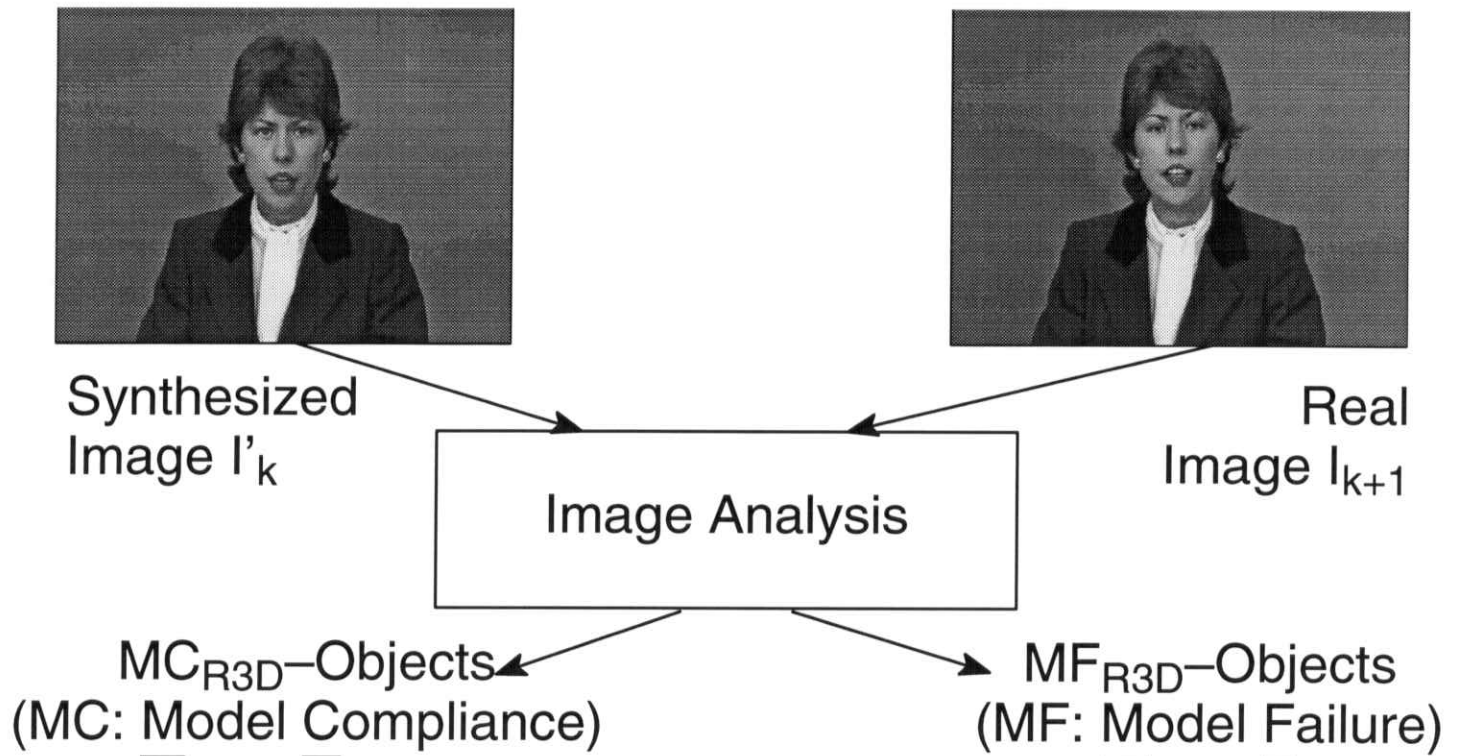


Areas of model failures

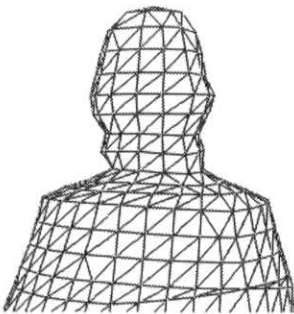


Geometrical distortions (Irrelevancy) are neglected

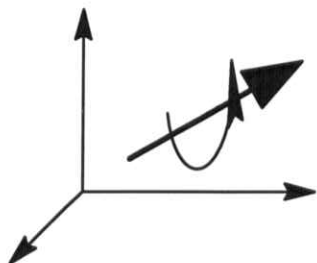
# OBASC: Image Analysis R3D



Texture

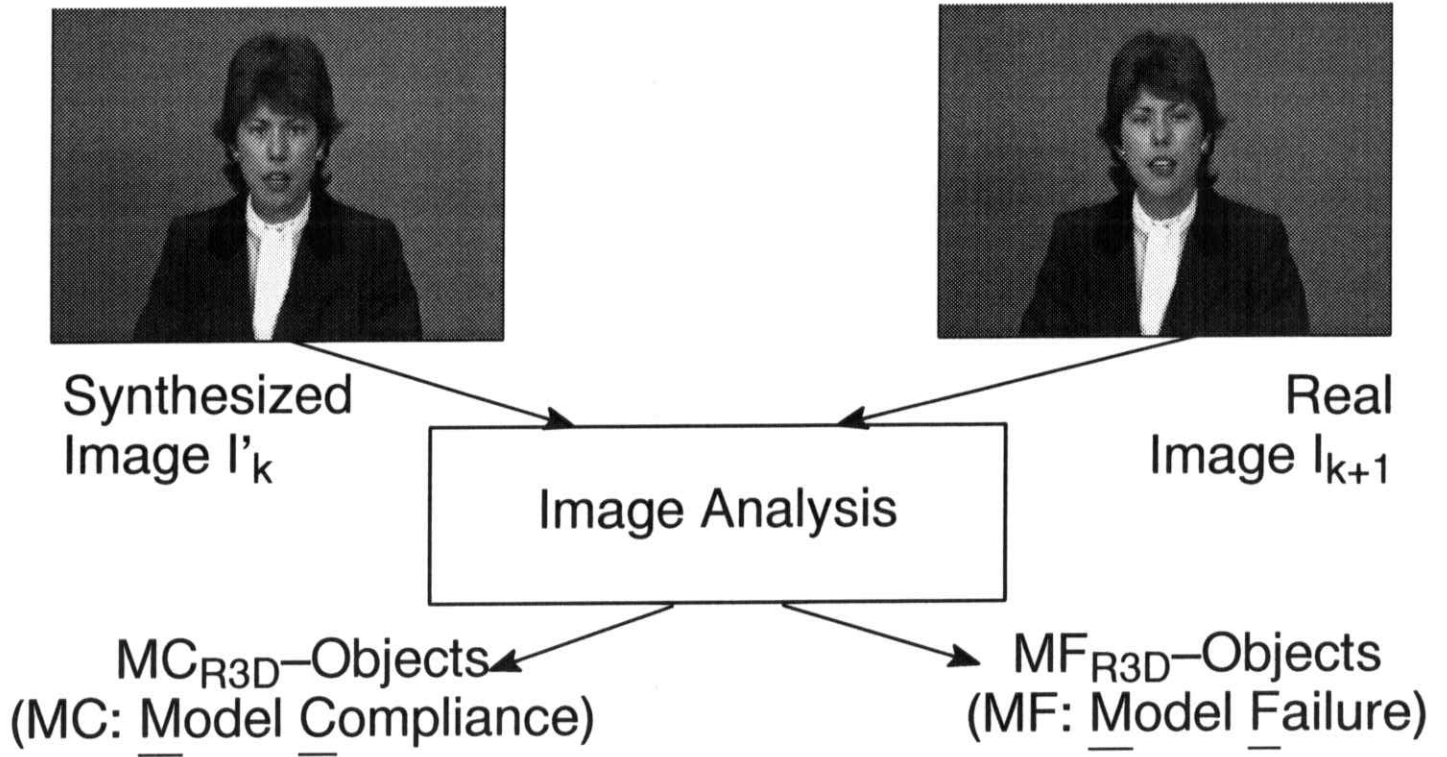


Shape



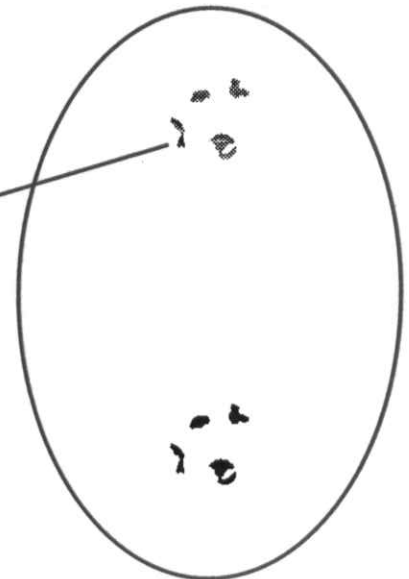
Motion

# OBASC: Parameter Coding R3D

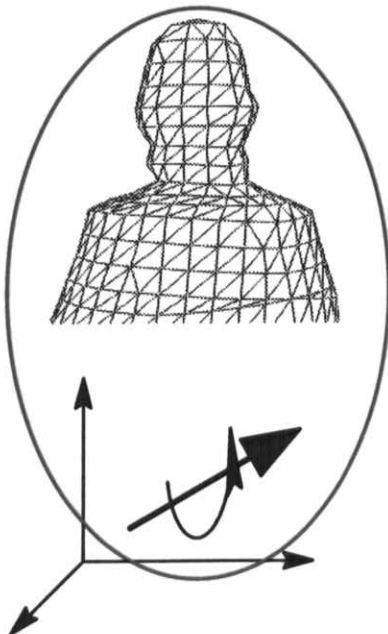


Texture

**1.1 bit/pel**  
**4% of image area**



Shape



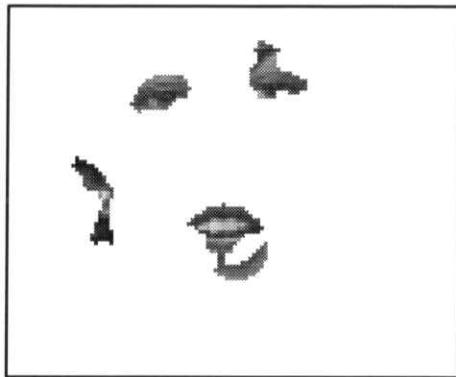
Motion

# OBASC: Image Analysis F3D

---



$MC_{R3D}$ -  
Objects



$MF_{R3D}$ -  
Objects



Real  
Image  $I_{k+1}$

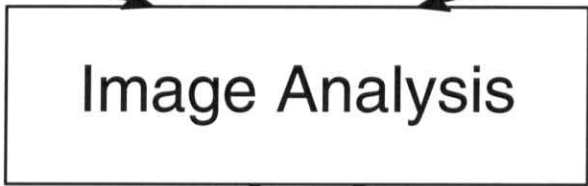


Image Analysis

$MC_{F3D}$ -Object

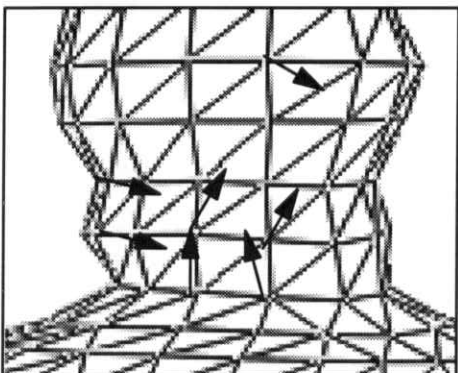
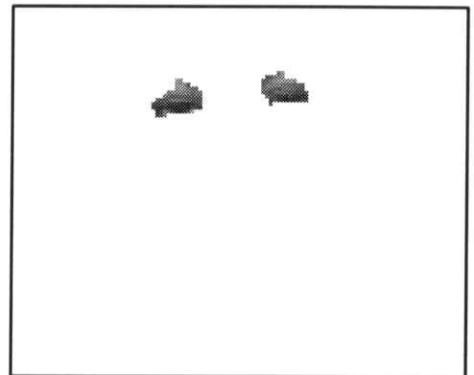
(MC: Model Compliance)

$MF_{F3D}$ -Objects

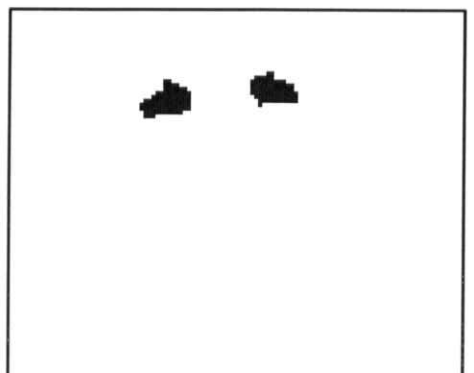
(MF: Model Failure)



Texture



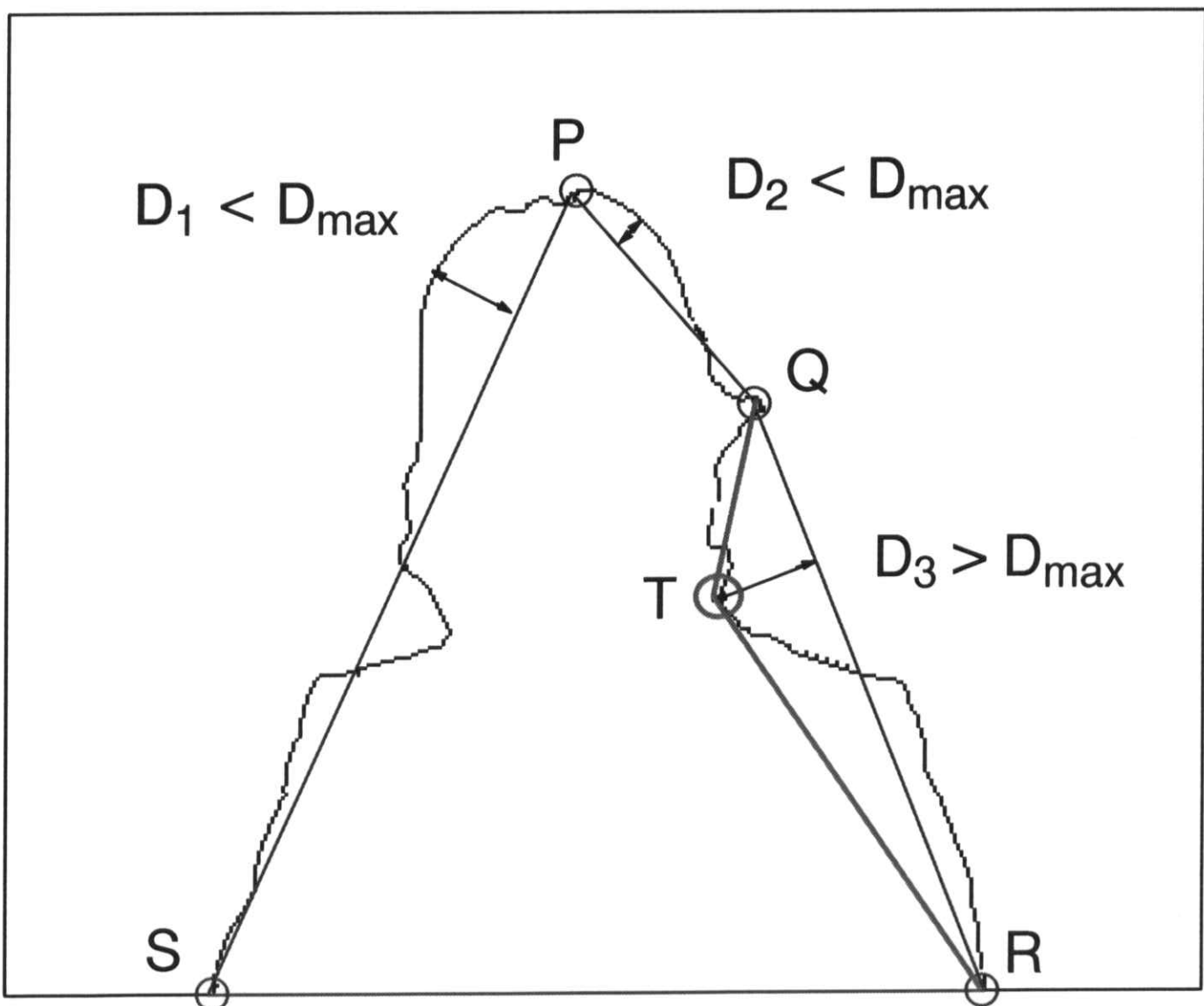
Shape



## 2D Shape Coding

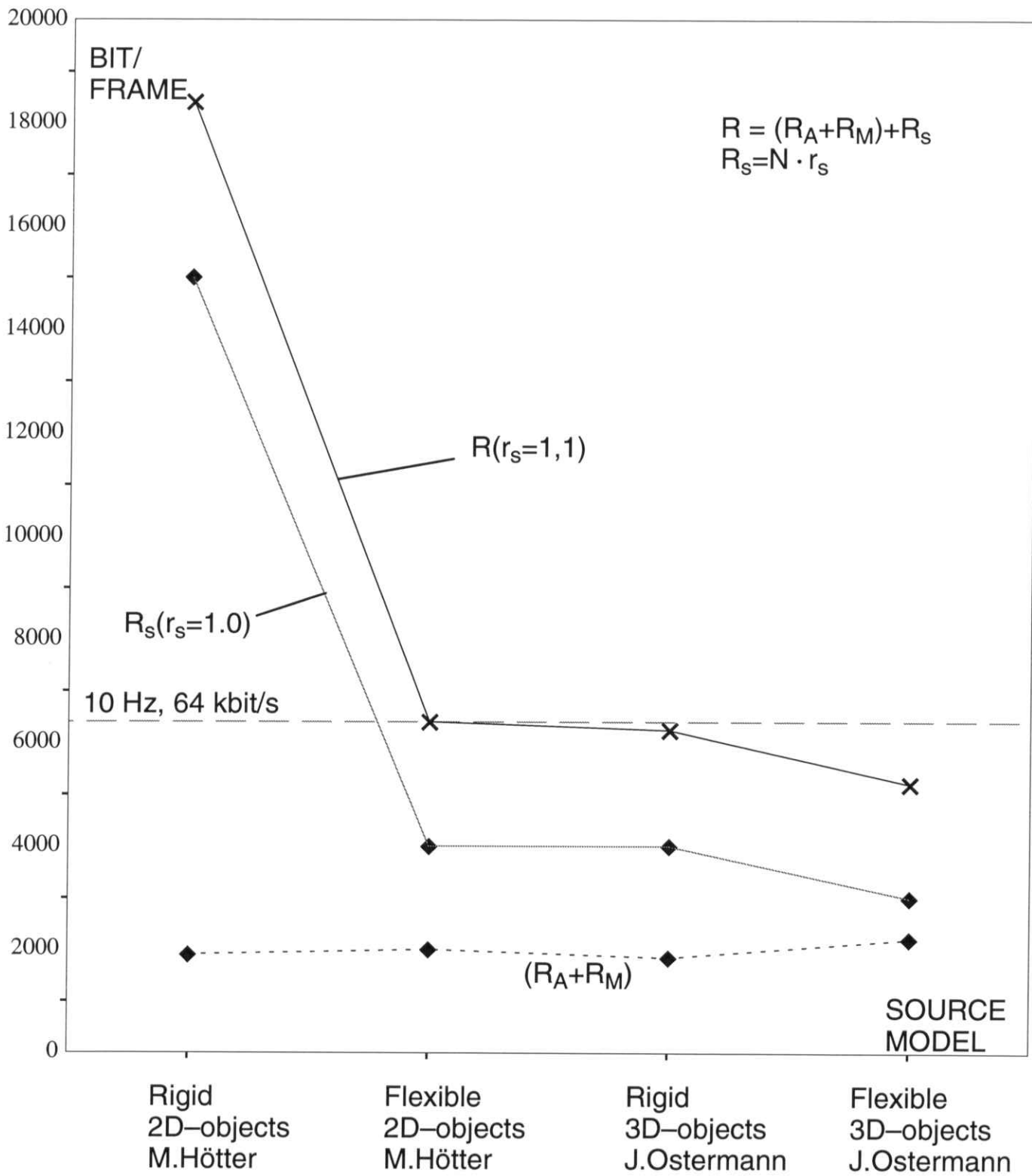
### Approximation of Object Shape by Polygons

- Maximum elongation  $\longrightarrow$  starting points (PQRS)
- Quality measure of the approximation :  
absolute distance between the original contour and the approximated contour :  $D \leq D_{\max}$



Improving Approximation with new Point T





**EXPERIMENTAL RESULTS OF AVERAGE BIT RATES  $R_A+R_M$  AND  $R$  IN BIT PER CIF-FRAME**

# Video Presentation

---

## Test Sequence *CLAIRE*

- Original
- Shape of Model Object
- MF–Objects
- Coded Video
  - Comparison F2D – R3D (64 kbit/s)
  - F3D (56 kbit/s)

SOURCE MODEL	ENCODED INFORMATION	CODING TECHNIQUE
PELS	COLOUR OF PELS	PCM
STATISTICALLY DEPENDENT PELS	COLOUR OF PELS OR BLOCK OF PELS	PREDICTIVE CODING TRANSFORM CODING
TRANSLATORICALLY MOVING BLOCKS	COLOUR OF BLOCKS AND MOTION VECTORS	MOTION COMPENSATED HYBRID CODING
MOVING UNKNOWN OBJECTS	SHAPE, MOTION AND COLOUR OF EACH OBJECT	OBJECT – BASED ANALYSIS – SYNTHESIS CODING
MOVING KNOWN OBJECT	SHAPE, MOTION AND COLOUR OF THE KNOWN OBJECT	KNOWLEDGE – BASED CODING
FACIAL EXPRESSIONS	ACTION UNITS	SEMANTIC CODING

# Knowledge-Based Coding: Extension of Source Model

---

## Problem:

- colour parameters of MF-objects require 60 % of data rate
- detection of MF-objects independent of image contents (object-based)

## Goal:

- reduction of MF-objects

## Approach:

- detection and tracking of faces
  - ◆ detection of eyes and mouth
  - ◆ adaptation of facial mask
  - ◆ face tracking with facial mask
- scene dependent detection of MF-objects (knowledge-based)

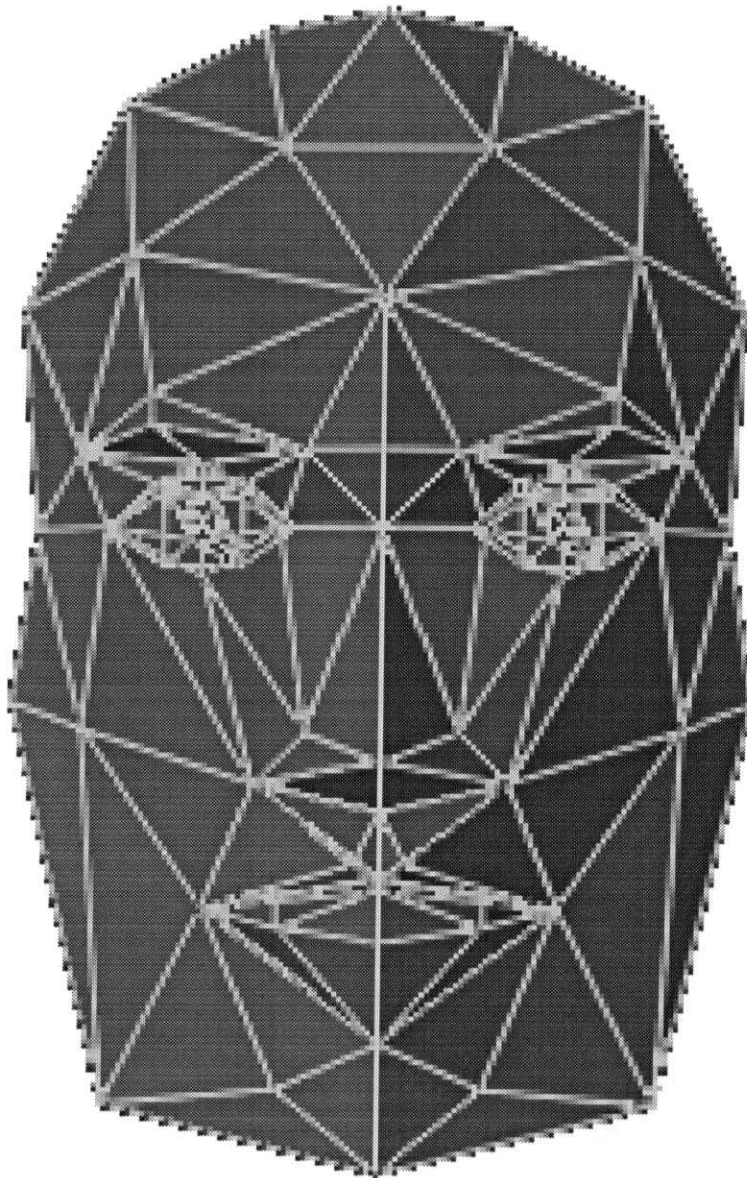
Outlook: semantic coding

# Knowledge-Based Coding: F3D with Candide

---

Additional shape parameters

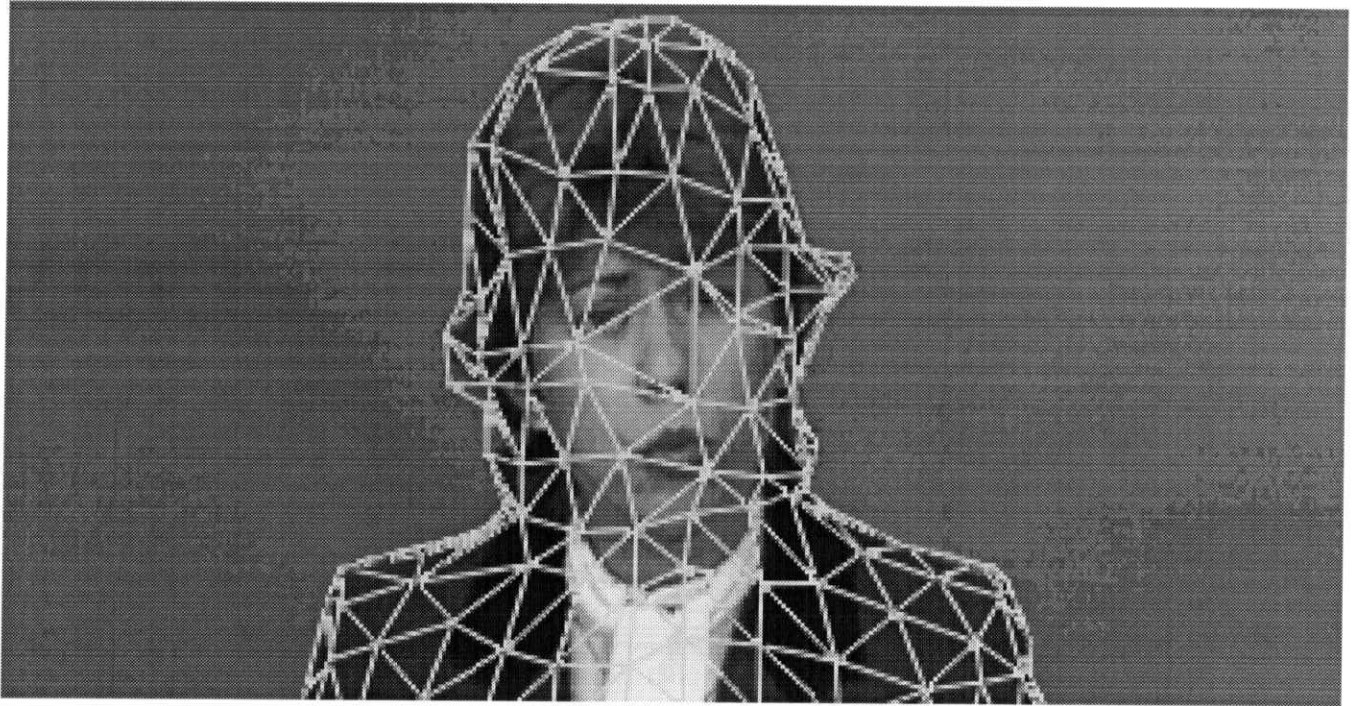
- mouth position
- eye positions



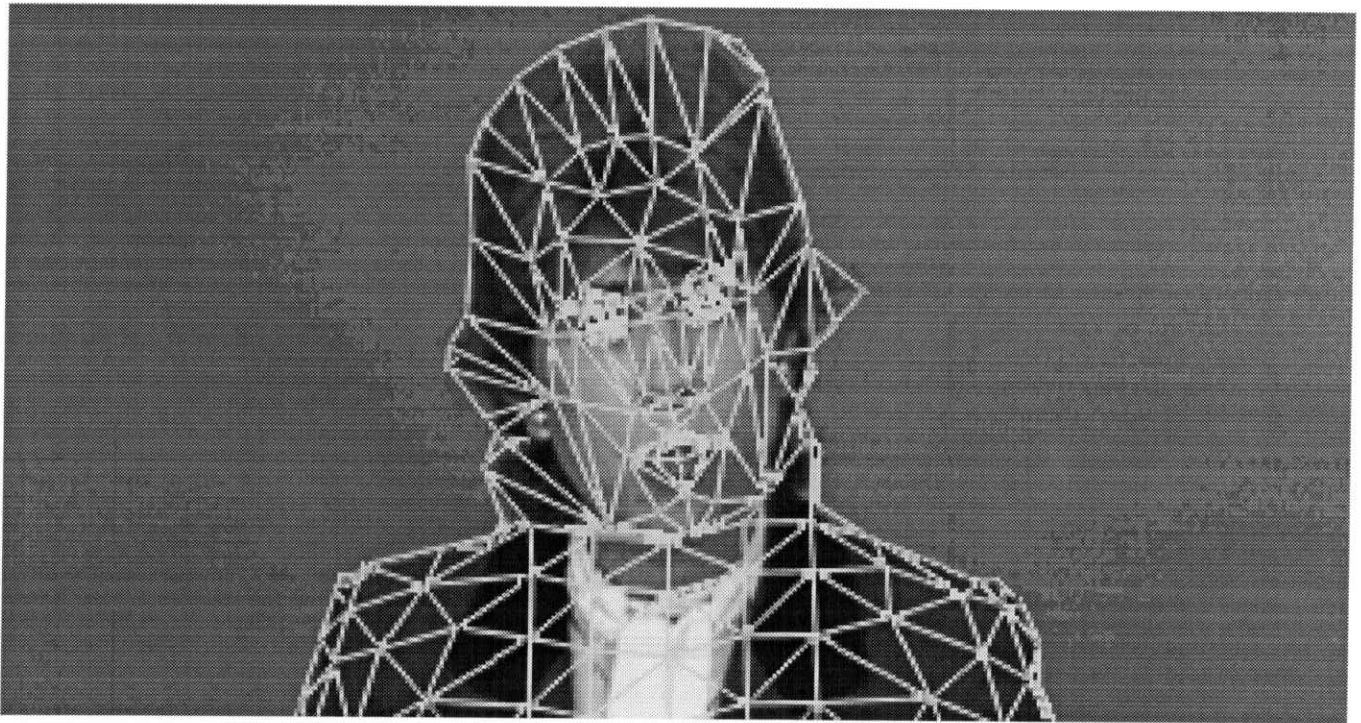
R. Rydfalk, "CANDIDE, A parameterised face", Internal Report Lith-  
ISY-I-0866, Linköping University, Linköping, Sweden, 1987

# Adaptation of Candide

---



model object F3D



model object F3D with Candide adapted

# Knowledge-Based Coding: Results

---

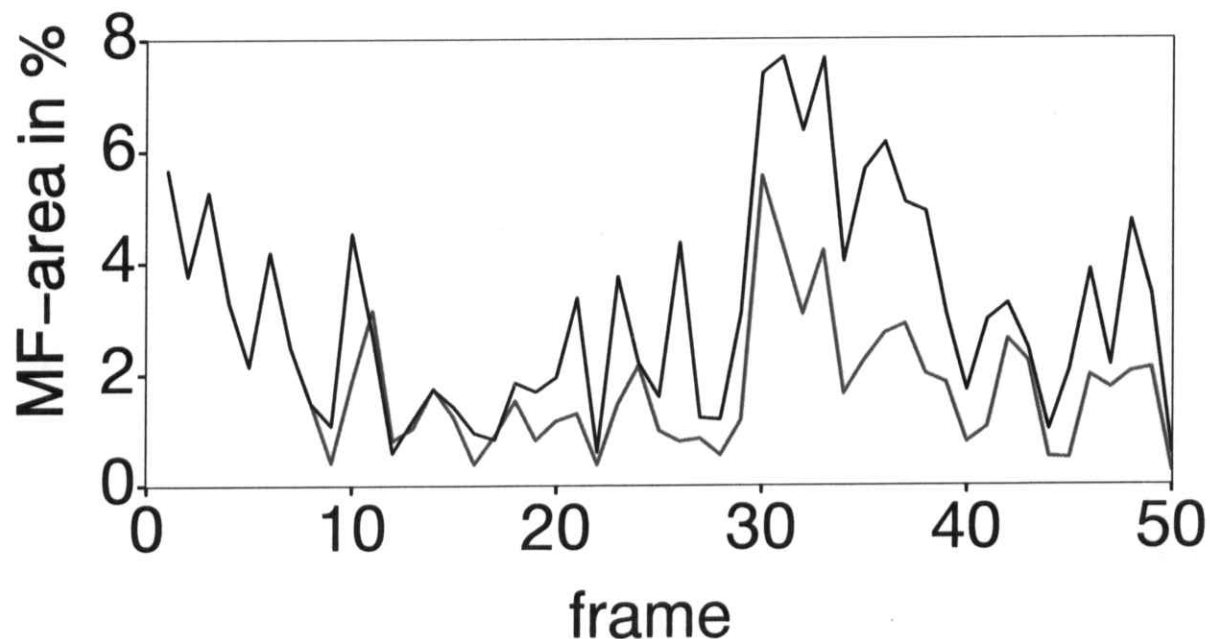
Number of frames to detect face

- Claire 9 frames
- Miss America 14 frames
- Michael 4 frames

Area of model failure in %

(object-based versus knowledge-based)

(Claire, CIF, 10 Hz)



# Conclusions

---

Source models for image coding

Block-based hybrid coding

- H.261, MPEG-1, MPEG-2, H.263

Object-based analysis-synthesis coding

- Image analysis
- Increased coding efficiency due to better modelling of the real world

Knowledge-based analysis-synthesis coding

- Increased coding efficiency due to scene knowledge
- Face model with limited degree of freedom

Computer graphics applied to image coding