Survey of Image and Video Quality Metrics based on Vision Models

Christian J. van den Branden Lambrecht
Hewlett-Packard Laboratories
1501 Page Mill Road MS 1U20
Palo Alto, CA 94304
vdb@hpl.hp.com

Outline

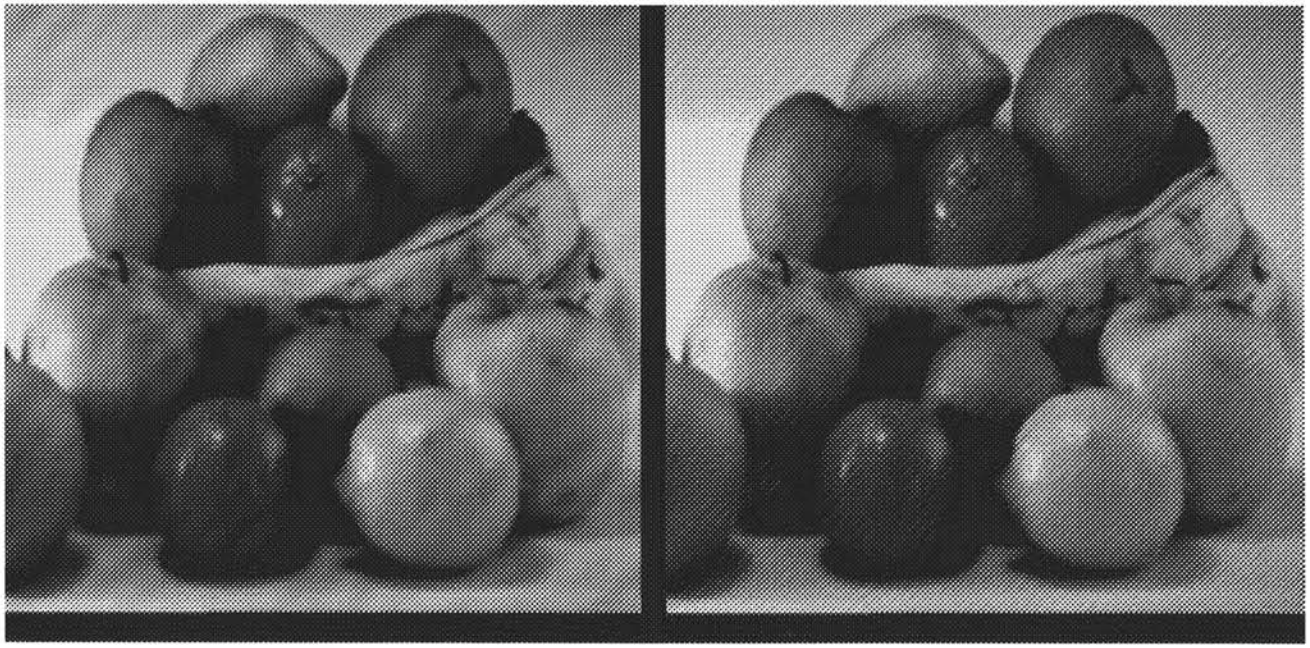
- **Why simple metric won't work.**
- **Review of related works.**
- **Principles of vision modeling.**
- **Examples:**
  - *model for still pictures,*
  - *model for video.*
- **Limitations of vision models.**
- **Conclusion.**

Why simple metric won't work

- **Perception depends on many factors:**
  - *viewing distance,*
  - *focus of attention,*
  - *contrast adaptation.*
- **Difficult to model these phenomena without *vision modeling*.**

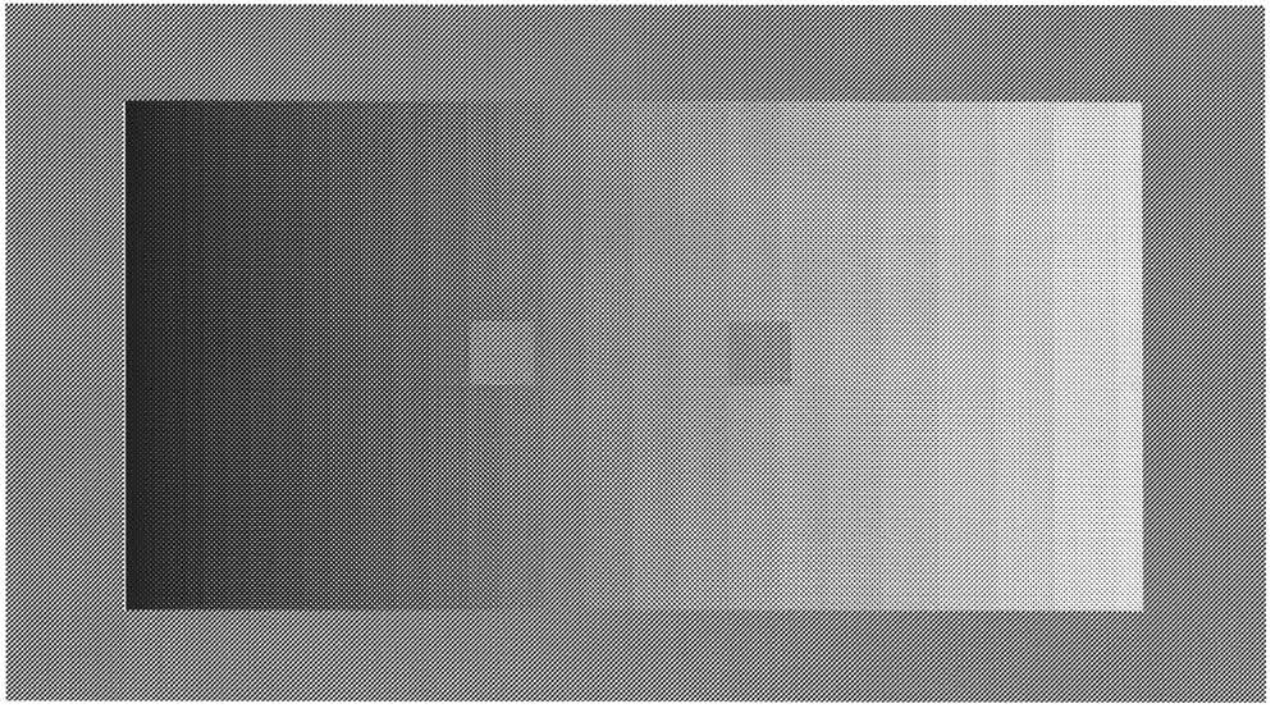
Contrast Sensitivity - Illustration

- **Both images have same amount of noise (i.e. same PSNR)**



Color Appearance: Example

- **Both patches have the same color:**



Related Works

- **Still pictures:**
  - *2D single-channel models*
    - *Schade.*
    - *Mannos & Sakrison.*
  - *Multi-channels models*
    - *Watson.*
    - *Daly.*
    - *Lubin*
    - *Comes.*
    - *Teo & Heeger.*

Related Works

- **Metrics for video:**
  - *Lukas & Budrikis.*
  - *van den Branden, Verscheure.*
  - *Osberger.*
  - *Western.*
  - *Lubin.*
  - *Bellofiore & Karam.*

Principles of Multi-Channel Models

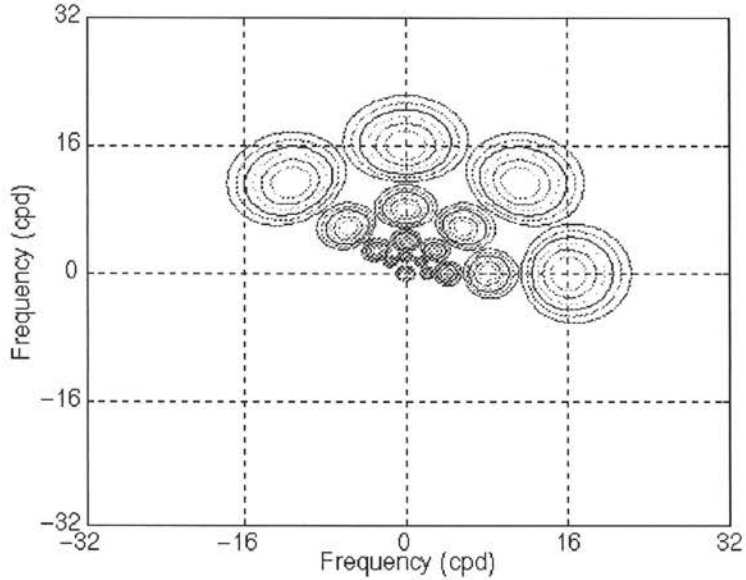
- **Existence of several visual streams:**
  - *High/low tempral frequencies, Color pathways.*
- **Information encoded by contrast:**
  - *Same perception of objects under very different illumination conditions.*
- **Representation of contrast at various scales and orientations.**
- **Sensitivity to contrast varies as a function of frequency, orientation and color.**
- **Stimuli interfere: phenomenon of visual masking.**

Multi-Resolution Architecture

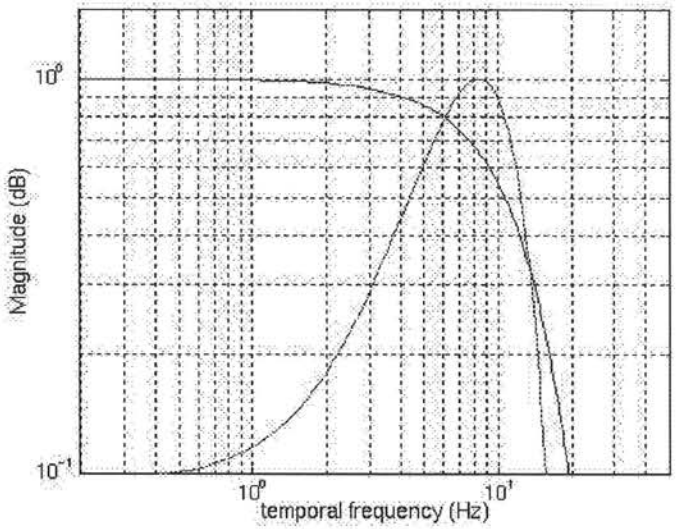
- **Cortical neurons exhibit a band-pass response.**
- **Existence of several mechanisms termed channels.**
- **Channels are tuned in:**
  - *spatial frequency,*
  - *orientation,*
  - *temporal frequency,*
  - *color.*

Illustration

- **Spatial Channels:**

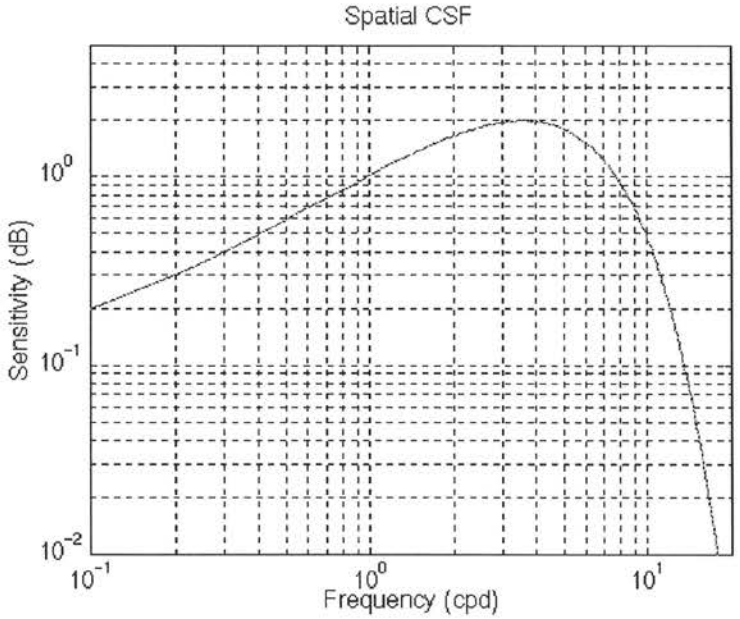


- **Temporal Channels:**

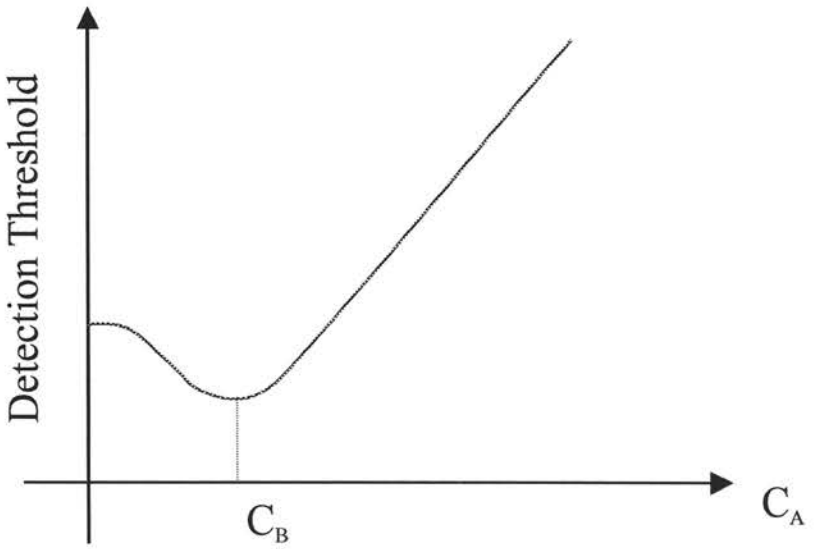


Pattern Sensitivity

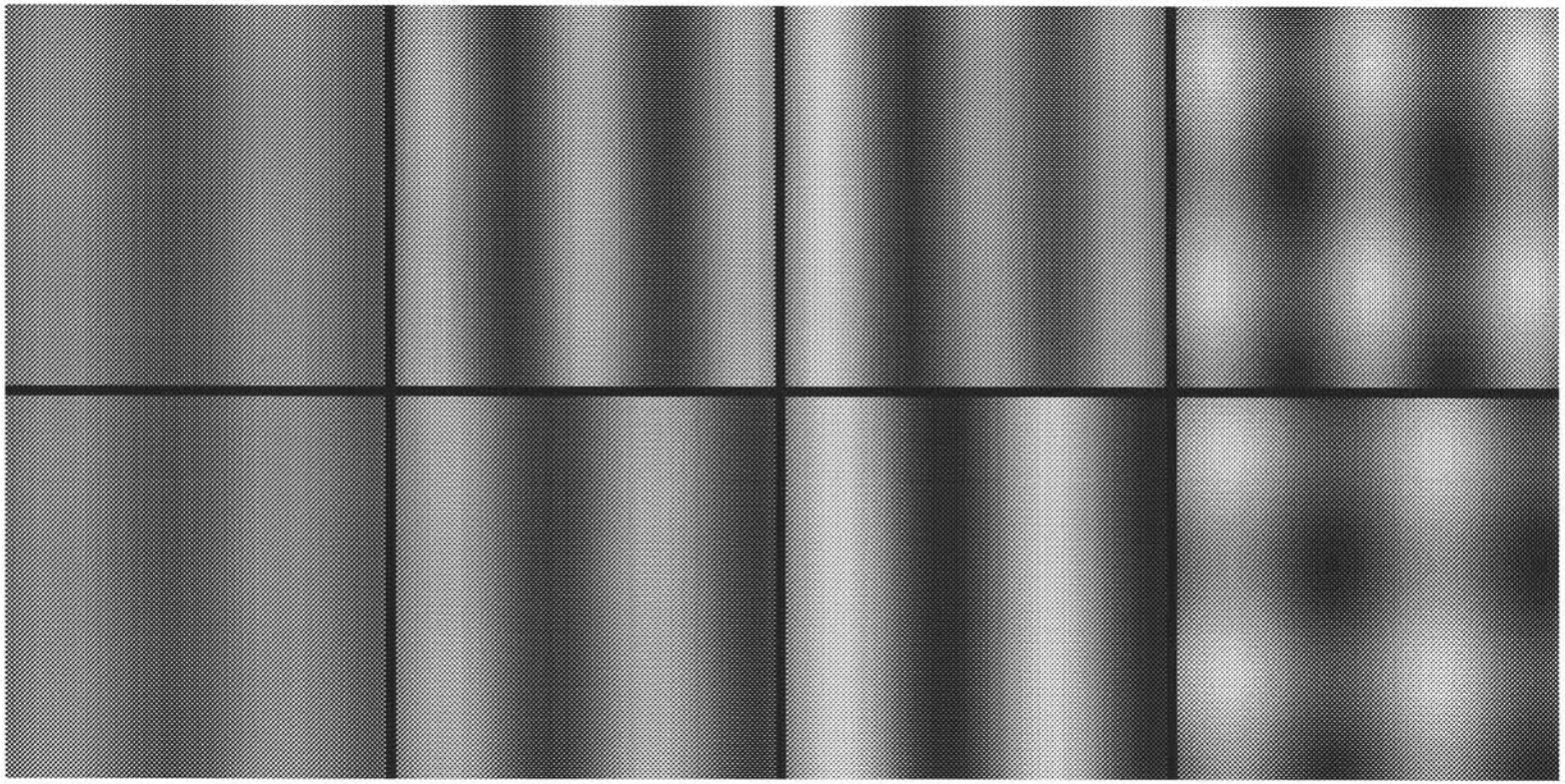
- **Contrast sensitivity:**
  - *function of the frequency, color and orientation*



- **Visual masking:**
  - *Modification of the detection threshold of a stimulus as a function of the local background contrast.*



Visual Masking - Illustration



Vision Models Parameterization

- **Parameters of sensory processes models have to be obtained by subjective experiments.**
- **Many methods:**
  - *Method of limit:*
    - *Subject obtain threshold stimulus by increasing from low stimulus and decreasing from high stimulus, then taking the mean.*
  - *Method of adjustment:*
    - *successive approximations oscillating around the threshold.*
  - *Staircase methods.*

Vision Models Parameterization

- **Experiments can be:**
  - *Detection,*
  - *Discrimination.*
- **Problem:**
  - *Can predict if a stimulus is visible or not.*
  - *No way to assess how visible.*

Teo & Heeger's Model

- **Models normalization of cortical receptive fields responses.**
- **Models inter-channel masking.**
- **More efficient implementation of the perceptual decomposition.**

Teo & Heeger's Model

- **Linear Transform**
- **Squaring**
- **Normalization**

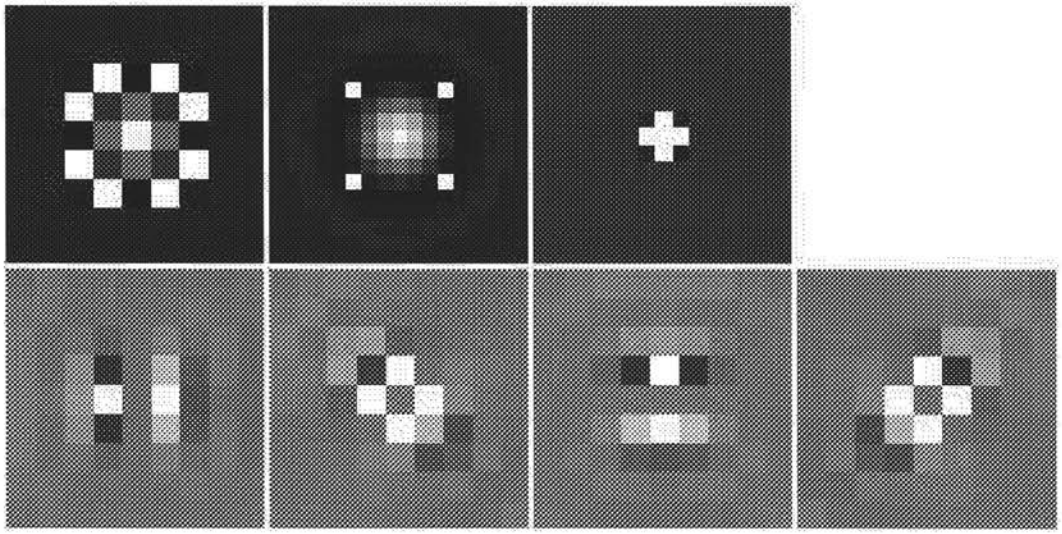
$$R^{\theta} = k \frac{\left(A^{\theta}\right)^2}{\sum_{\varphi} \left(A^{\varphi}\right)^2 + \sigma^2}$$

- **Detection**

$$\Delta R = \left| R_o - R_d \right|^2$$

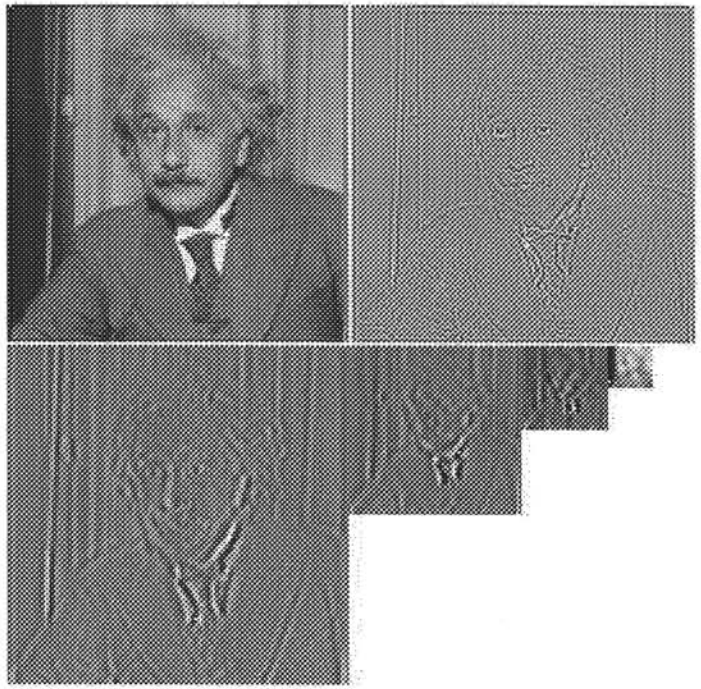
Teo & Heeger's Model

- **Linear Transform:**
  - *Steerable pyramid from Simoncelli* et al.



Teo & Heeger's Model

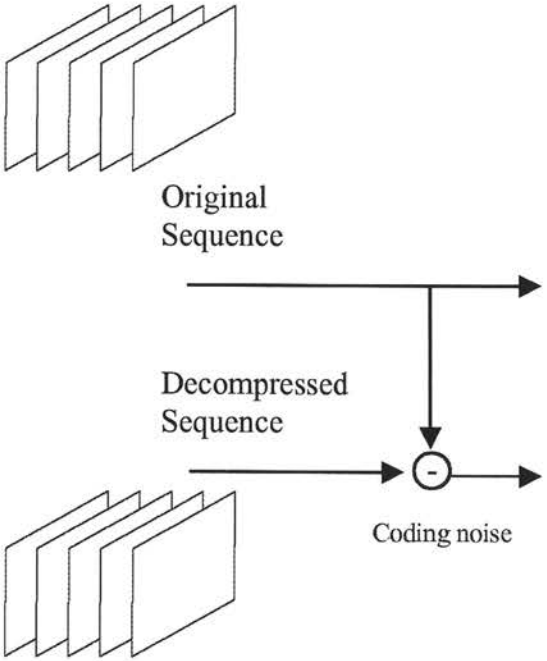
- **Linear Transform:**
  - *Steerable pyramid from Simoncelli* et al.



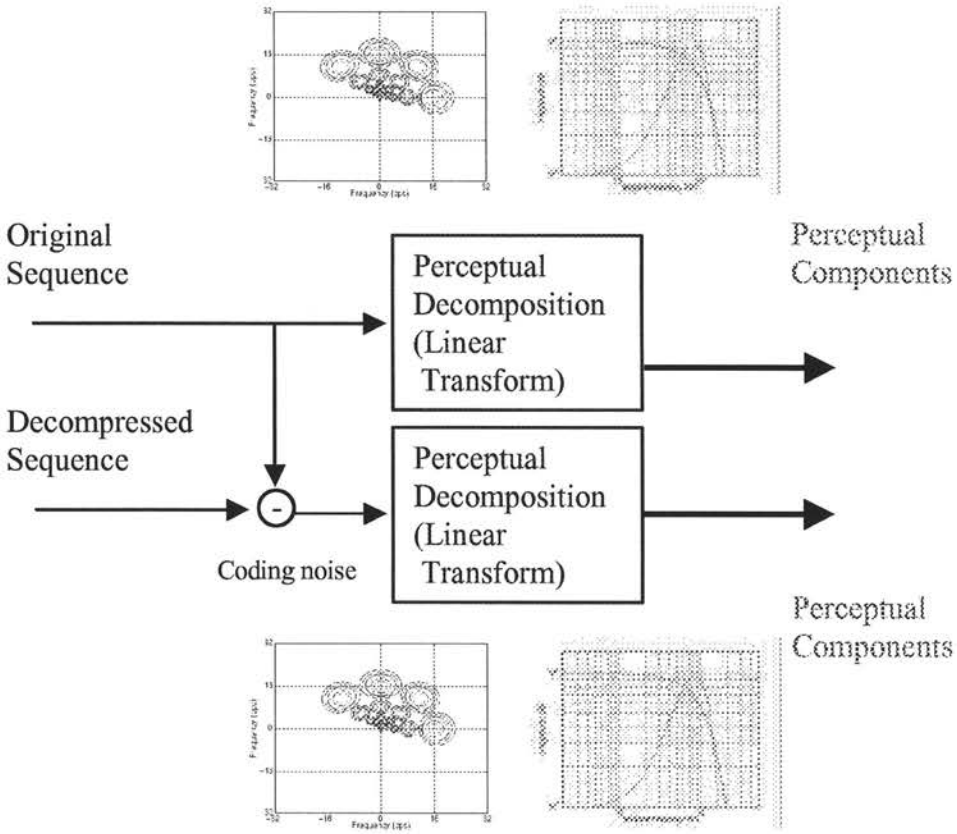
Vision Model for Video

- Several implementations.
- Multi-channel.
- Intra & inter channel masking.
- Color processing.
- Measured CSF wrt model of coding noise.

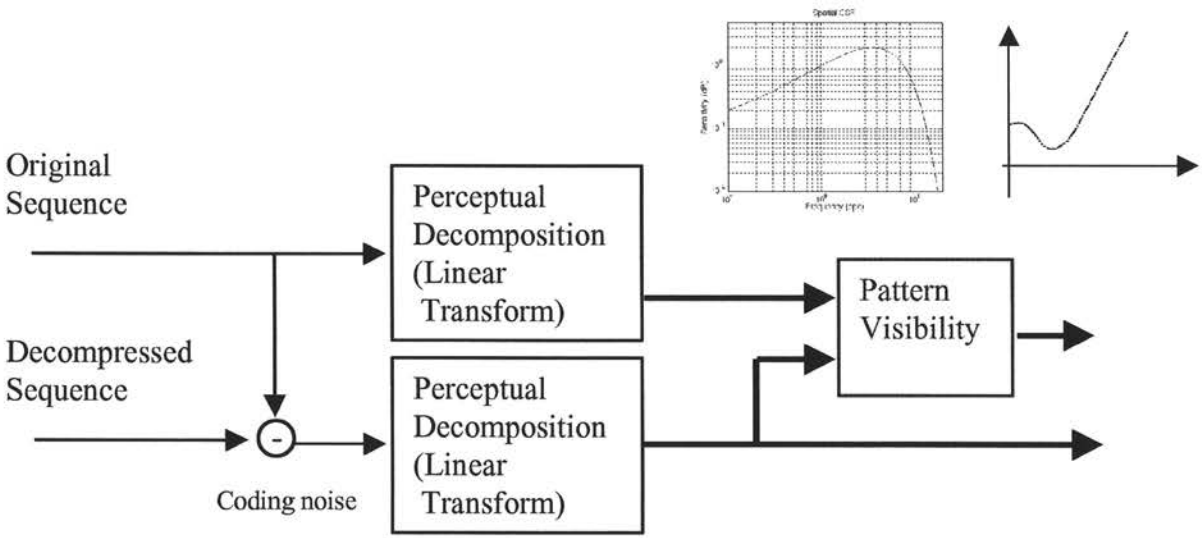
General Architecture



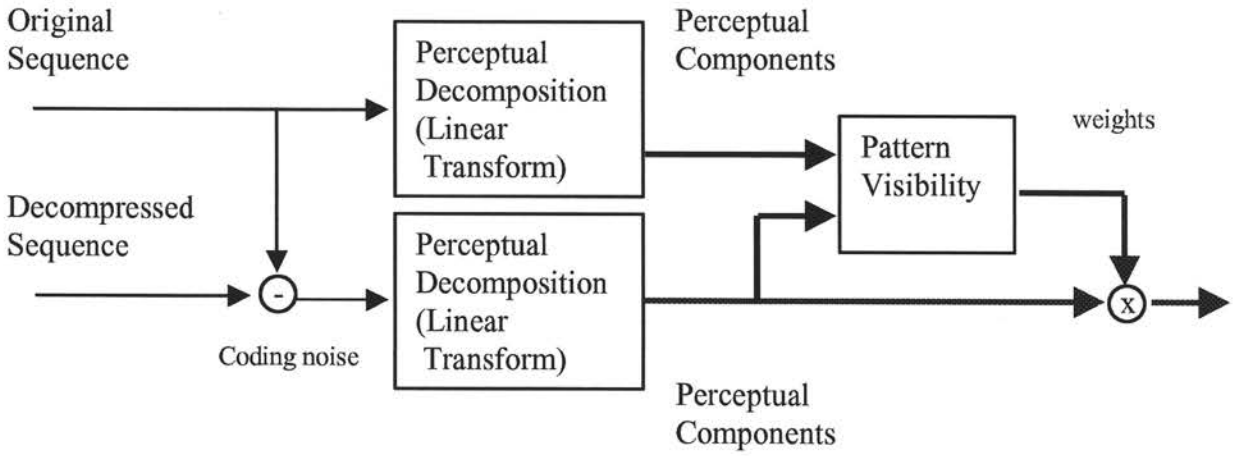
General Architecture



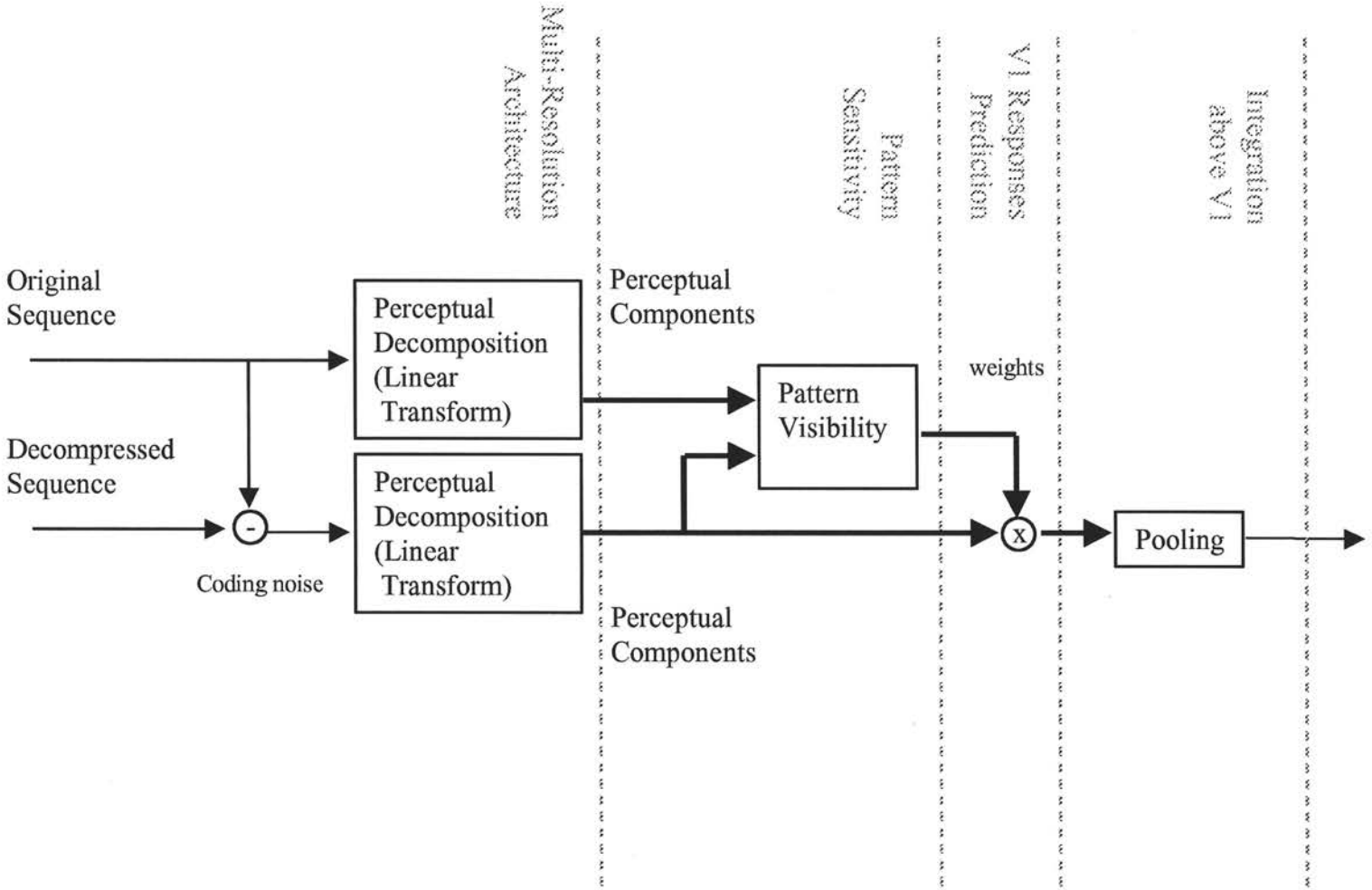
General Architecture



General Architecture



General Architecture



Exple: Moving Pictures Quality Metric

- **Computes the non-masked coding noise.**
- **Pools the data over channels to account for higher level of cognition.**
- **Computes a distortion measure on blocks of the sequence to account for:**
  - *persistence on the retina,*
  - *focus of attention.*

Output of the Metric

- **Logarithmic scale**
  - *Similar to decibels*
  - *Data expressed in "visual decibels"*

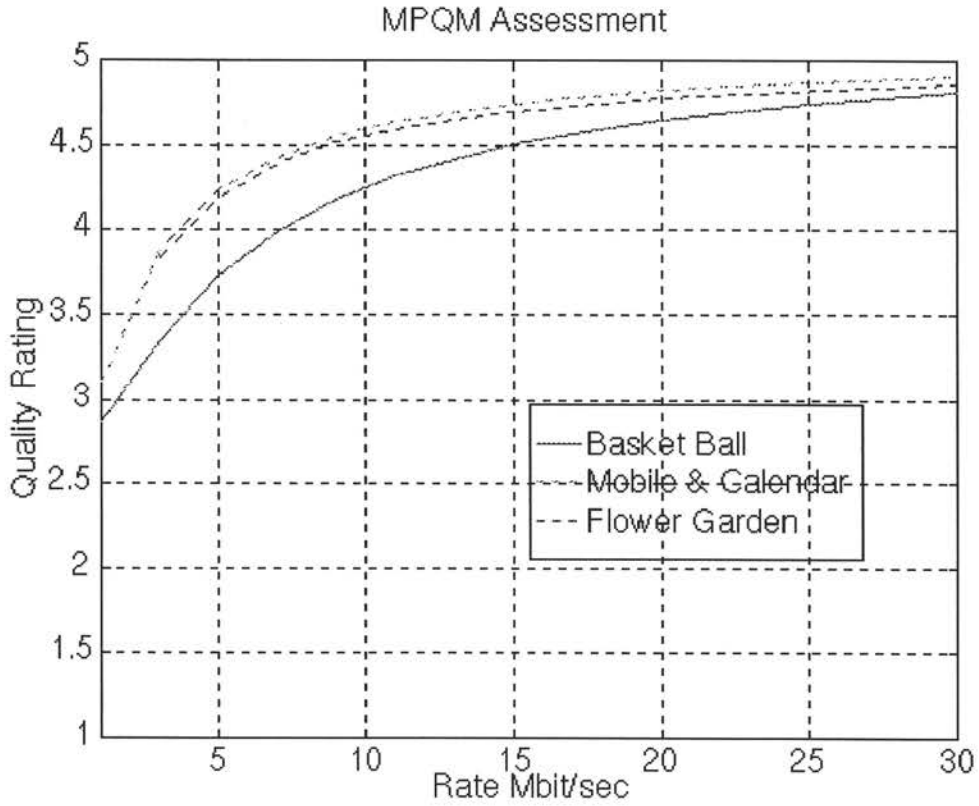
$$10\log_{10}\frac{255^2}{E^2}$$

- **CCIR 500 1 to 5 scale**
  - *Scaling factor chosen on the basis of the vision model.*

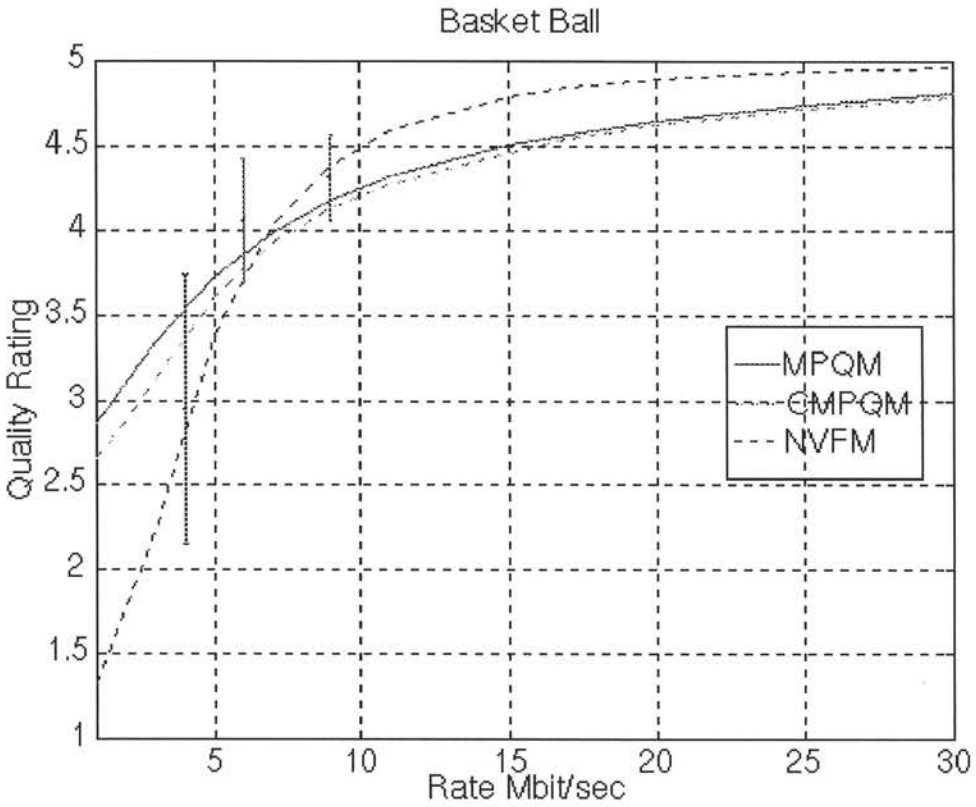
$$Q=\frac{5}{1+NE}$$

Rating	Impairment	Quality
5	Imperceptible	Excellent
4	Perceptible	Good
3	Slightly annoying	Fair
2	Annoying	Poor
1	Very annoying	Bad

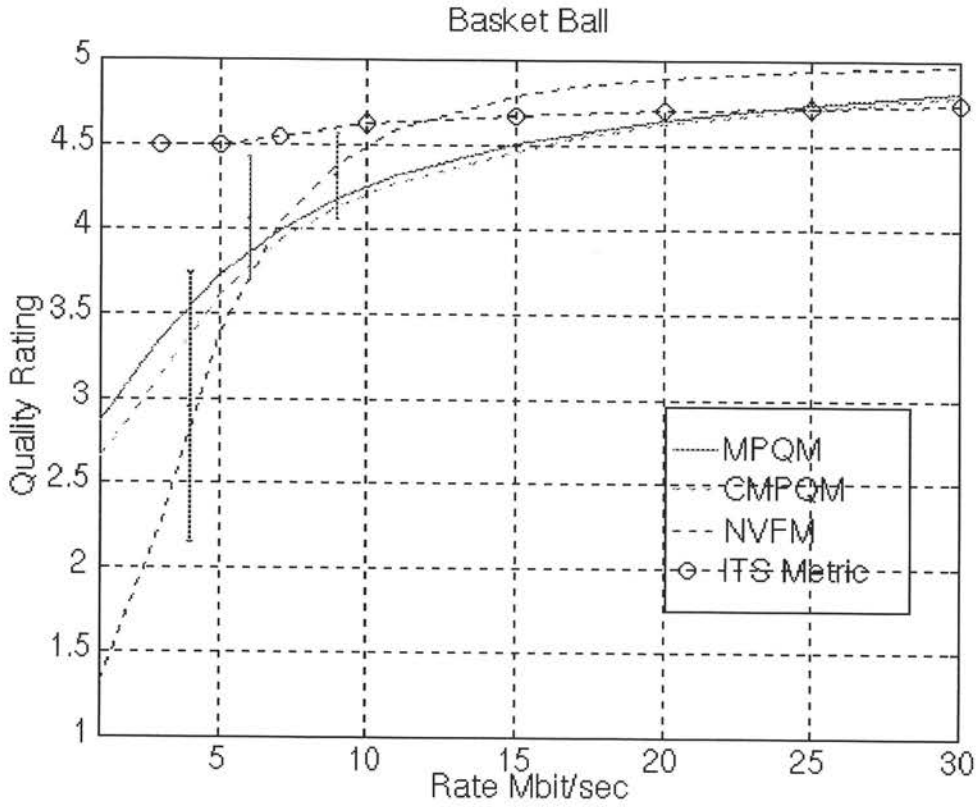
Results


MPQM Assessment
Quality Rating
Rate Mbit/sec
5
4.5
4
3.5
3
2.5
2
1.5
1
5
10
15
20
25
30
Basket Ball
Mobile & Calendar
Flower Garden

Comparison with Subjective Data



Comparison with Subjective Data



Extension of the previous model

- Perfect Reconstruction filter bank.
- Color Extension (based on opponent-colors theory),
- Normalization model.
- Motion model.
- Metrics for image features.

Limitations of Vision Models

- **Many parameters, still some unknowns about the visual system.**
- **Can be accurate around threshold of vision:**
  - *bearly visible artifacts*
- **Cannot assess the visibility of highly distorted pictures.**

Conclusion

- **A good quality metrics needs vision modeling:**
  - *accounts for many subjective aspects.*
- **Vision models are not mature yet:**
  - *limited in use,*
  - *not stable, need large scale validation.*
- **Common effort of the industry and academia could lead to better metrics:**
  - *Discussion& experiments could be done in the public domain (T1A1?).*