Visual perception of human motion

Niko Troje
BioMotionLab

Department of Psychology
& School of Computing
Queen's University
www.BioMotionLab.ca

Outline

- Vision from the psychologist's point of view: A bit of history and a few concepts
- Biological motion: perception and analysis

Perception is …

… to infer a consistent, reliable, predictive model of the world given noisy, incomplete, ambiguous sensory data!

?

Stimulus | Sensation | Perception

Gestalt psychology

Gestalt (German): form, shape, figure

Kurt Koffka
(1886 – 1941)

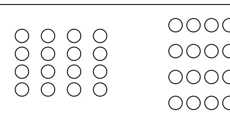
Wolfgang Köhler
(1887 – 1967)

Max Wertheimer
(1880 – 1943)

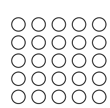
The whole is more than the sum of its parts!

Gestalt "Laws"

Law of proximity



Law of similarity



Law of good continuation

Law of common fate

Gestalt theory

- Law of pragnanz
  - Prägnanz: German for good figure, design, carrying meaning
  - A stimulus pattern is organized in such a way that the resulting structure is as simple as possible.
  - "The Simplicity Principle" (e.g. Pomerantz & Kubovy 1986)

Simplicity

Stimulus — Sensation — Perception

Simplicity Principle

- Simplicity
  - Occam's razor (William of Ockham, 14th century): "entities must not be multiplied beyond necessity" (entia non sunt multiplicanda praeter necessitatem)
  - Information theory, channel capacity, redundancy (Shannon and Weaver 1949)
  - Coding theory (Hochberg & McAlister 1953, Leeuwenberg 1969, 1971; Restle 1979)
  - Attneave 1954, 1981: "What the perceptual system likes is short descriptions"
  - Minimum description length (Rissanen 1989, Hinton & Zemel 1994)

Hermann von Helmholtz (1821 – 1894)

- Measured the speed of a nerve impulse
- Invented the ophthalmoscope
- Important publications
  - "Handbook of Physiological Optics"
  - "The Sensation of Tone as a Physiological Basis for the Theory of Music"
  - "On the Conservation of Force"

Very good reading is a speech he gave in 1878 "The Facts of Perception"
http://www.marxists.org/reference/subject/philosophy/works/ge/helmholt.htm

Hermann von Helmholtz (1821 – 1894)

- Helmholtz advocated the "Likelihood Principle"
  - Perception as "unconscious inference"
  - Sensory input will be organized into the most probable distal object or event consistent with that input

Statistics

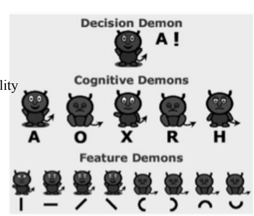
Stimulus — Sensation — Perception

Likelihood Principle

- Likelihood
  - Helmholtz's "unconscious inference"
  - Oliver Selfridge's (1959) "pandemonium" model of letter recognition. Committee rules: Honor physics, avoid accidents
  - Bayesian statistical decision theory:

$$p(w_i|s) \approx p(s|w_i) * p(w_i)$$

posterior probability — likelihood — prior probability "prior"

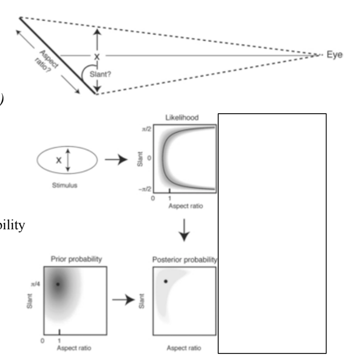


Bayes'ian Inference

$$p(w_i|s) \approx p(s|w_i) * p(w_i)$$

posterior probability — likelihood — prior probability

$$C = \arg\max_{w_i} [p(w_i|s)]$$

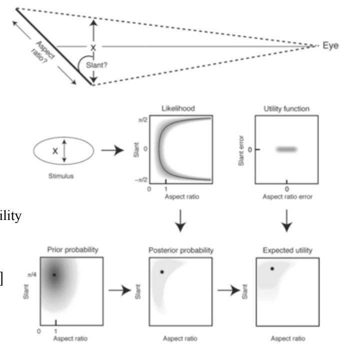


Bayes'ian Inference

$p(w_i|s) \approx p(s|w_i) * p(w_i)$

posterior probability — likelihood — prior probability

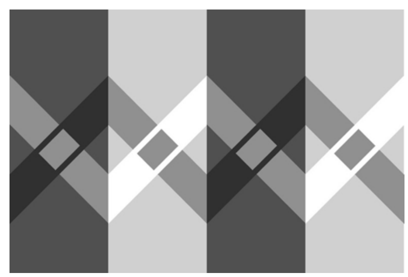
$R = \arg\max_r [u(r,W)\, p(W|s)]$



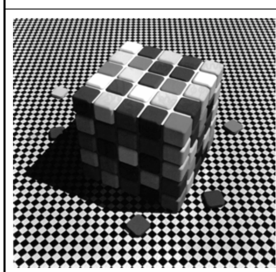
More examples for ambiguities in visual perception

what we want to know		what we can measure	context
aspect ratio	slant (of an ellipse)	aspect ratio of the retinal image	our example
reflectance	illuminance	luminance	lightness constancy
size	distance	retinal size	size – distance ambiguity
intensity	wavelength	absorption	principle of univariance
object motion	motion of the retina	retinal motion	efference copy theory
speed	direction of motion	image motion	aperture problem
3D shape	direction of light	2D retinal image	light-from-above bias
distance	motion of observer relative to object	retinal motion	motion parallax

Criss-cross Illusion



E. H. Adelson, 2000, www-bcs.mit.edu/gaz/



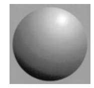
luminance = illuminance * reflectance

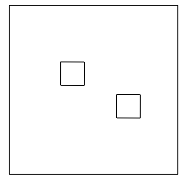
More examples for ambiguities in visual perception

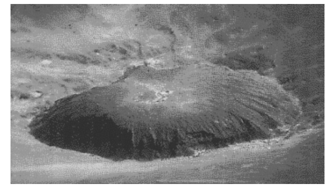
what we want to know		what we can measure	context
aspect ratio	slant (of an ellipse)	aspect ratio of the retinal image	our example
reflectance	illuminance	luminance	lightness constancy
size	distance	retinal size	size – distance ambiguity
intensity	wavelength	absorption	principle of univariance
object motion	motion of the retina	retinal motion	efference copy theory
speed	direction of motion	image motion	aperture problem
3D shape	direction of light	2D retinal image	light-from-above bias
distance	motion of observer relative to object	retinal motion	motion parallax

Light-from-above bias

Light-from-above bias







Leeuwenberg & Boselie 1988

Simplicity vs Likelihood

- Simplicity or Likelihood as an organizing principle?
  - What is role of redunancy?
  - Simplicity: "entities must not be multiplied beyond necessity" (Occam's razor)
  - Likelihood: Correlations in the data are necessary to make predictions

Simplicity = Likelihood

- It can be shown that there is no conflict and that the principles are in fact equivalent
  - Cover & Thomas 1991, Rissanen 1989, Mumford 1992
  - Chater 1996, Kolmogorov 1965)

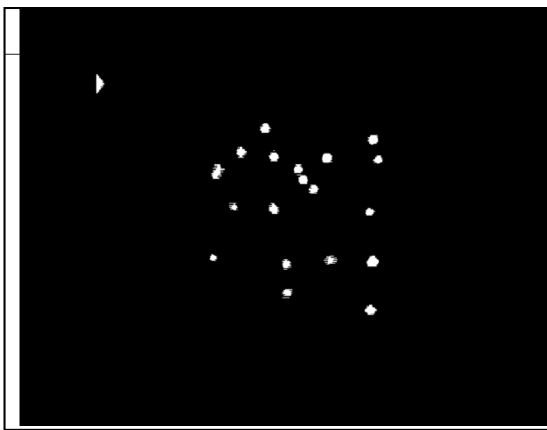
Horace Barlow (1921 -)

- What is efficient coding?
  - Efficient coding should convert hidden redundancy into a manifest, explicit, immediately recognizable form, rather than reduce it or eliminate it.

Part II

- Biological motion: perception and analysis

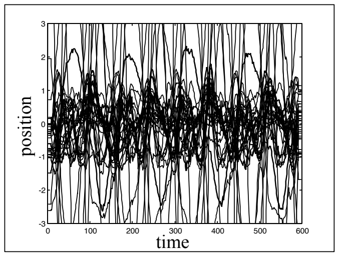
Gunnar Johansson
1911 - 1995



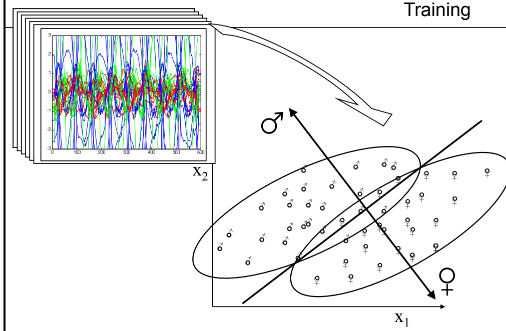
Starting point

- Biological motion contains plenty of information
- Our visual system is well able to retrieve and interpret this information

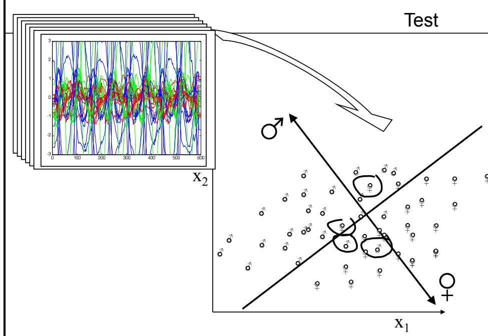
Question	How can we retrieve information encoded in animate motion?
Example	What makes a walker appear male or female?
Principles	A more general view
Applications	… in computer vision, animation, gait analysis



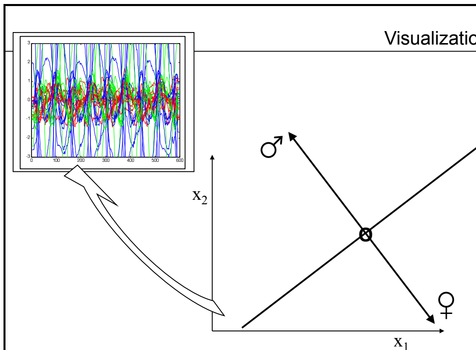
Training

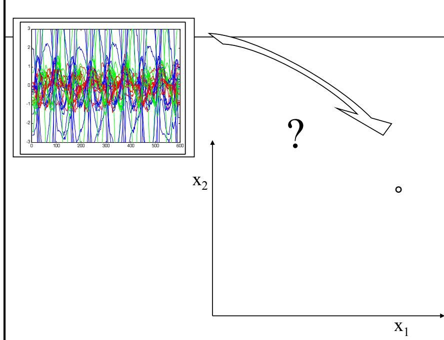


Test



Visualization





The goal is to find a representation that
- is loss-less
- supports linear operations
- is low-dimensional

"morphable model"

The approach:
- Start with a time series of postures expressed as Cartesian coordinates in a body centred coordinate system
- Express each trajectory as a discrete, low-dimensional Fourier series

The rest is simple:
- Reduce dimensionality by means of principle components analysis on the set of walkers
- Linear regression

$$p(t) = p_0 + a_1 \sin(\omega t + \varphi_1) + a_2 \sin(2\omega t + \varphi_2) + \ldots$$

$$p(t) = p_0 + p_1 \sin(\omega t) + q_1 \cos(\omega t) + p_2 \sin(2\omega t) + q_2 \cos(2\omega t) + \ldots$$

k	power [%]
1	94.2
2	4.4
3-n	1.4

$$p(t) = p_0 + a_1 \sin(\omega t + \varphi_1) + a_2 \sin(2\omega t + \varphi_2) + \ldots$$

$$p(t) = p_0 + p_1 \sin(\omega t) + q_1 \cos(\omega t) + p_2 \sin(2\omega t) + q_2 \cos(2\omega t) + \ldots$$

$w = [P, \omega]$

$$P = \begin{pmatrix} p0_{1,x} & p1_{1,x} & q1_{1,x} & p2_{1,x} & q2_{1,x} \\ p0_{2,x} & p1_{2,x} & q1_{2,x} & p2_{2,x} & q2_{2,x} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p0_{1,y} & p1_{1,y} & q1_{1,y} & p2_{1,y} & q2_{1,y} \\ p0_{2,y} & p1_{2,y} & q1_{2,y} & p2_{2,y} & q2_{2,y} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p0_{1,z} & p1_{1,z} & q1_{1,z} & p2_{1,z} & q2_{1,z} \\ p0_{2,z} & p1_{2,z} & q1_{2,z} & p2_{2,z} & q2_{2,z} \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

It's a morphable representation!

$$w = \begin{bmatrix} p0_{1,x} \\ p0_{2,x} \\ p0_{3,x} \\ p0_{4,x} \\ \vdots \\ p0_{n,z} \\ p1_{1,x} \\ \vdots \\ q2_{n-1,z} \\ q2_{n,z} \\ \omega \end{bmatrix}$$

$$\frac{w_1 + w_2}{2}$$

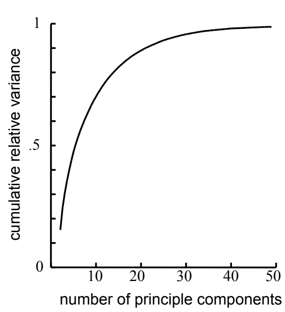
$$\frac{\Sigma\, w_i}{n}$$

$$\frac{\Sigma\, (a_i\, w_i)}{\Sigma\, a_i}$$

$$w = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \\ p_m \end{bmatrix}$$

$$W = \begin{pmatrix} p_{1,1} & p_{1,2} & & p_{1,k} \\ p_{2,1} & p_{2,2} & & p_{2,k} \\ p_{3,1} & p_{3,2} & & p_{3,k} \\ \vdots & \vdots & \cdots & \vdots \\ p_{m,1} & p_{m,2} & & p_{m,k} \end{pmatrix}$$

Principle component analysis



training set / test data

$$W = \begin{pmatrix} p_{1,1} & p_{1,2} & & p_{1,k} \\ p_{2,1} & p_{2,2} & & p_{2,k} \\ p_{3,1} & p_{3,2} & & p_{3,k} \\ \vdots & \vdots & \cdots & \vdots \\ p_{m,1} & p_{m,2} & & p_{m,k} \end{pmatrix} \qquad t = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ \vdots \\ ? \end{pmatrix}$$

$W_d = M\, W_p$ $\qquad$ $t_d = M\, t_p$

W_p: predictors $\qquad$ W_d: dependent variables training set | test data

$$W = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,k} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,k} \\ p_{3,1} & p_{3,2} & \cdots & p_{3,k} \\ & \ldots & & \\ p_{m,1} & p_{m,2} & \cdots & p_{m,k} \end{pmatrix} \qquad t = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ \vdots \\ ? \end{pmatrix}$$

$$W_d = M\, W_p \qquad t_d = M\, t_p$$

☐ W_p: predictors ☐ W_d: dependent variables

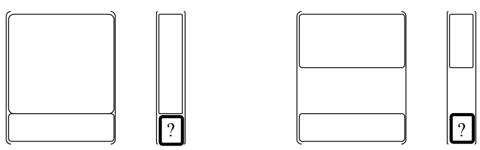
Example: gender classification

- Data: 50 male and 50 female individuals
  Over-ground walking, 15 landmarks (mostly joints)
  Fourier-based morphable model
- Take one pattern out and use the rest to learn the model
  Apply the model to the remaining pattern
  Evaluate the outcome
  Repeat the procedure for every pattern
- Result: 92.5% correct classification!

Troje JOV 2002

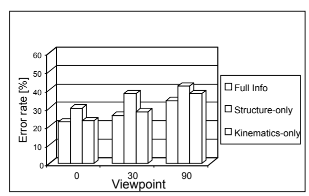
use all information (3D, kinematics & shape) to predict gender

use only parts (2D) info to predict gender

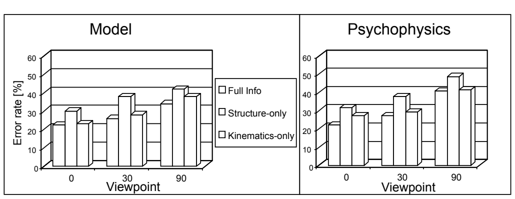


☐ predictors ☐ dependent variables

Classification performance of the model



Classification performance of the model and of human observers



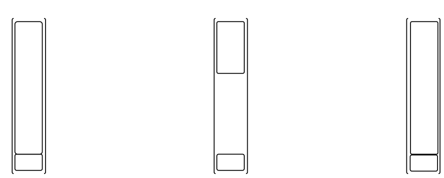
24 observers, 3 s presentations (PLD)
View: 0, 30, 90 deg (between subject)
Info: full info, structure-only, dynamic-only (within subject)

Troje JOV 2002

use all information (3D, kinematics & shape) to predict gender

use only parts (e.g. 2D) info to predict gender

use gender to predict kinematics



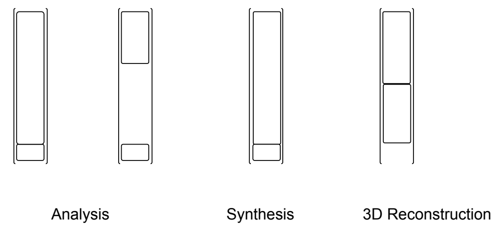
Analysis

Synthesis

☐ predictors ☐ dependent variables

Gender was just an example.
Use the same idea for other attributes:

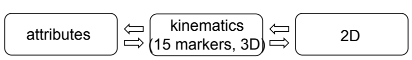
- biological: weight, size, age
- emotional
- personality traits
- clinical: psychology, neurology, biomechanics



Zhang, Z., and Troje, N. F. (2007) 3D Periodic Human Motion Reconstruction from 2D Motion Sequences. Neural Computation 19:1400-1421.

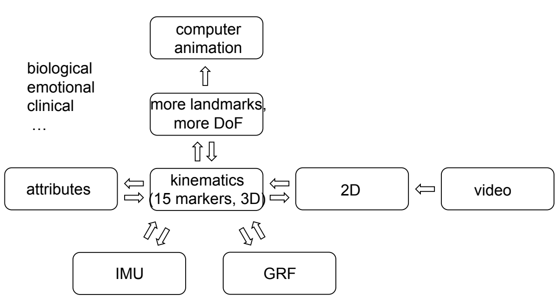
Zhang, Z., and Troje, N. F. (2005) View-independent person identification from human gait. Neurocomputing 69:250-256.

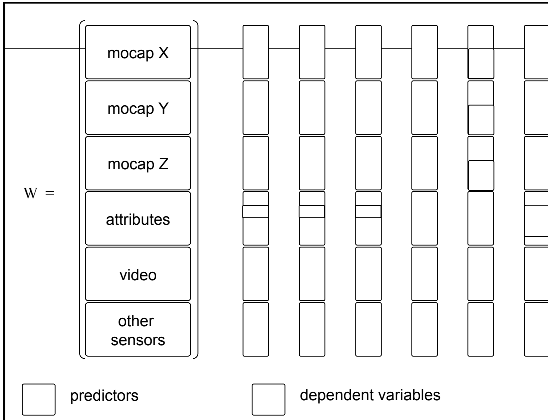
biological
emotional
clinical
…



Sigal, L., Troje, N. F., Fleet, D. (2010) Human Attributes from 3D Pose Tracking Search Results. 11th European Conference on Computer Vision (ECCV).

biological
emotional
clinical
…





The goal is to find a representation that

- supports linear operations
- is loss-less
- is low-dimensional

"morphable model"

The approach:

- Start with a time series of postures expressed as Cartesian coordinates in a body centred coordinate system
- Express each trajectory as a discrete, low-dimensional Fourier series

The rest is simple:

- Reduce dimensionality by means of principle components analysis on the set of walkers
- Linear pattern classification

The human visual system is a proof of concept and a source of inspiration for how to infer information from noisy, incomplete, ambiguous sensory data!

Journal of Cognitive Neuroscience, 1991

Eigenfaces for Recognition

Matthew Turk and Alex Pentland
Vision and Modeling Group
The Media Laboratory
Massachusetts Institute of Technology

Abstract

■ We have developed a near-real-time computer system that can locate and track a subject's head, and then recognize the person by comparing characteristics of the face to those of known individuals. The computational approach taken in this system is motivated by both physiology and information theory, as well as by the practical requirements of near-real-time performance and accuracy. Our approach treats the face recognition problem as an intrinsically two-dimensional (2-D) recognition problem rather than requiring recovery of three-dimensional geometry, taking advantage of the fact that faces are normally upright and thus may be described by a small set of 2-D characteristic views. The system functions by projecting face images onto a feature space that spans the significant variations among known face images. The significant features are known as "eigenfaces," because they are the eigenvectors (principal components) of the set of faces; they do not necessarily correspond to features such as eyes, ears, and noses. The projection operation characterizes an individual face by a weighted sum of the eigenface features, and so to recognize a particular face it is necessary only to compare these weights to those of known individuals. Some particular advantages of our approach are that it provides for the ability to learn and later recognize new faces in an unsupervised manner, and that it is easy to implement using a neural network architecture. ■

FIAS

Vetter & Troje, 1997

---

published in: Journal of the Optical Society of America A 14:9 (1997) 2152-2161.

Separation of texture and shape in images of faces for image coding and synthesis

THOMAS VETTER AND NIKOLAUS F. TROJE
vetter|troje@mpik-tueb.mpg.de
Max-Planck-Institut für Biologische Kybernetik, Spemannstr. 38, 72076 Tübingen, Germany

Abstract. Human faces differ in shape and texture. Image representations based on such a separation have been reported by several authors [for a review, see Beymer and Poggio, (1996)]. This paper investigates such a representation of human faces based on a separation of texture and two-dimensional shape information. Texture and shape were separated using pixel-by-pixel correspondence between the different images, which was established through algorithms known from optical flow computation. The paper demonstrates the improvement of the proposed representation over well established pixel-based techniques in terms of coding efficiency and in terms of the ability to generalize to new images of faces. The evaluation is performed by calculating different distance measures between the original image and its reconstruction and by measuring the time human subjects need to discriminate them.

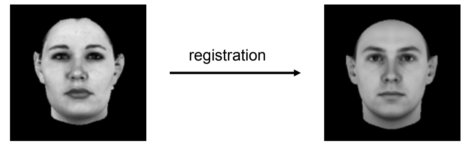
Keywords: Image synthesis, face recognition, flexible templates

Pixel-based representation

Correspondence-based representation – morphable model

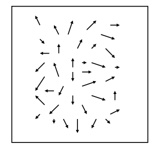
Vetter & Troje, 1997

Correspondence-based representation



"texture"

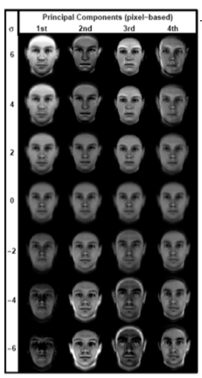
$$T = \begin{pmatrix} I_{1,1} \\ I_{1,2} \\ \vdots \\ I_{n,m} \end{pmatrix}$$



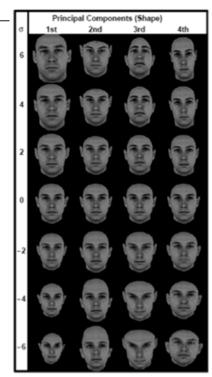
"shape"

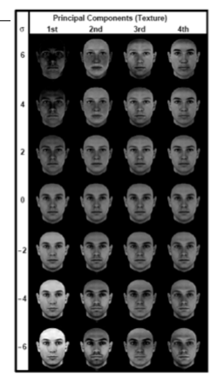
$$S = \begin{pmatrix} dx_{1,1} \\ dy_{1,1} \\ dx_{1,2} \\ \vdots \\ dx_{n,m} \\ dy_{n,m} \end{pmatrix}$$

Vetter & Troje, 1997

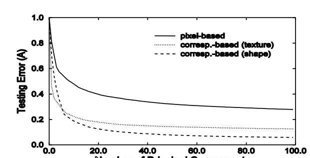


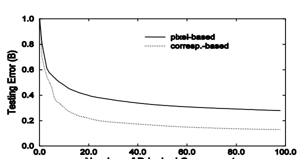
Vetter & Troje, 1997





Vetter & Troje, 1997





Vetter & Troje, 1997

Blanz & Vetter, 1999