

Computer Assisted Visual InterActive Recognition (CAVIAR) Technology

Arthur Evans, John Sikorski, Patricia Thomas, Sung-Hyuk Cha, Charles Tappert

Pace University, White Plains, NY

Jie Zou, Abishek Gattani, George Nagy

RPI DocLab, Troy, NY

nagy@ecse.rpi.edu

5/18/2005

Nagy IEEE EIT May 2005

1

Human and computer abilities differ

***** HUMANS :

Gestalt perception for object/background segmentation

Broad context

Correlated (colored) noise

Judgment of pairwise similarity

***** COMPUTERS :

Store millions of image-label pairs

Calculate geometrical moments, spatial frequencies, connected components, multivariate probabilities, decision rules, grammars

5/18/2005

Nagy IEEE EIT May 2005

2

To exploit these differences, we have built two experimental systems: CAVIAR-flower and CAVIAR-face

COMPUTER Builds a *visible model* (segments the object)

Corrects model (if needed)

COMPUTER Extracts features based on the model
Ranks candidates

HUMAN

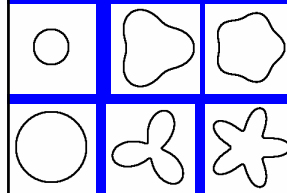
Decides final classification

5/18/2005

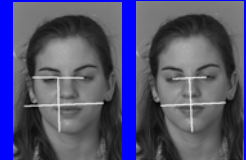
Nagy IEEE EIT May 2005

3

EXAMPLES OF VISIBLE MODELS



five characteristic points



rose curves

quadratic curves



5/18/2005

Nagy IEEE EIT May 2005

THE VISIBLE MODEL

- Mediates between human and computer.
- Domain-specific (different for flowers, faces, fruit, ...).
- Constructed by the computer.
Corrected by user if necessary .
- The model guides feature extraction.
The features are used for **rank ordering** the classes.
The reference pictures of the top candidates are displayed.
- The operator selects the reference picture most like the unknown picture.
- **The human is always in charge.**

5/18/2005

Nagy IEEE EIT May 2005

5

CAVIAR-flower GUI designed for outlining petals

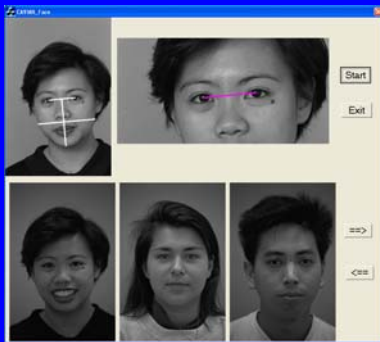


5/18/2005

Nagy IEEE EIT May 2005

6

CAVIAR-face GUI designed for accurate pupil location

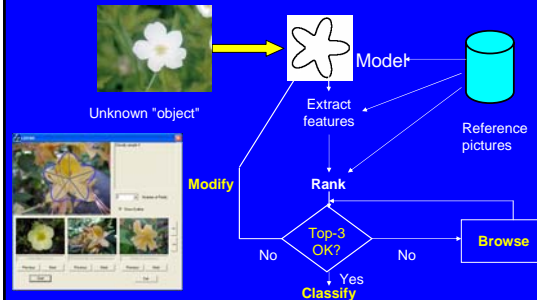


5/18/2005

Nagy IEEE EIT May 2005

7

CAVIAR DATA FLOW



5/18/2005

Nagy IEEE EIT May 2005

8

CAVIAR-FLOWER COMPARED TO MACHINE ALONE AND TO HUMAN ALONE.

102 classes, 102 unknowns, 6 subjects

	Accuracy (%)	Time per flower (seconds)
Interactive	93 (83 - 99)	12 (7 - 27)
Machine Alone	32 (24 - 50)	-
Human Alone	93 (91 - 97)	26 (18 - 36)

5/18/2005

Nagy IEEE EIT May 2005

9

CAVIAR-FACE COMPARED TO MACHINE ALONE AND TO HUMAN ALONE (200 faces)

200 pictures as gallery, 50 pictures as probes, 6 subjects

	Accuracy (%)	Time per face (seconds)
Interactive	99.7	7.6
Machine alone	47.0	--
Human alone	--	66.3

5/18/2005

Nagy IEEE EIT May 2005

10

CONCLUSIONS

Interactive recognition is twice as **fast** as unaided human, and twice as **accurate** as unaided machine (without years of R&D).

Parsimonious interaction *throughout* the process is better than only at the beginning or end.

Interactive systems can be initialized with a *single training sample per class*, and *improve with use*.

Interaction with images requires a *visible model* that is accessible to both man and machine.

Let both do what they do best.

Leave the human in charge.

5/18/2005

Nagy IEEE EIT May 2005

11

SELF-CONTAINED MOBILE CAVIAR AT PACE UNIVERSITY



Sharp Zaurus



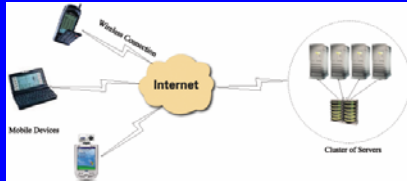
5/18/2005

Nagy IEEE EIT May 2005

12

M-CAVIAR

- CAVIAR implemented on mobile devices.
- Able to classify patterns in situ.
- Client-server Architecture



5/18/2005

Nagy IEEE EIT May 2005

13

MOBILE CAVIAR AT RENSSELAER



5/18/2005

Toshiba

Nagy IEEE EIT May 2005

Abhishek Gattani

14

OTHER APPLICATIONS: FISH ??



Black Crappie



Alabama Shad



Atlantic Sturgeon



Blue Gill

U.S. Fish & wild life service

5/18/2005

Nagy IEEE EIT May 2005

15

CRYPTIC CATS ?



Jan Schipper
NSF-IGERT Fellow
CATIE
Escuela Posgrado
Sede Central 7170
Turrialba, Costa Rica
Central America

Proyecto Conservación del Área Talamanca (ProCAT) is an international project under the umbrella of the Institute of the Rockies.

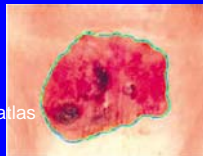
5/18/2005

Nagy IEEE EIT May 2005

16

CAVIAR-Derma?

- Nearly 1000 diagnoses (classes)
- Big image atlases available
 - John Hopkins dermatology image atlas
 - University of Erlangen



For: eczemas, scar assessment, insect/plant/frost bite, burns, STD, spots (measles, chickenpox), cosmetics, bioterror (cutaneous anthrax, tularemia, smallpox; military, expeditions, isolated elderly, developing countries, training paramedics, ...

5/18/2005

Nagy IEEE EIT May 2005

17

MARIGOLDS



Digital camera

Nikon Coolpix 775



PDA

Veo 130s



Cell phone

Motorola V400

5/18/2005

Nagy IEEE EIT May 2005

18

Thank you!

www.ecse.rpi.edu/doclab/EIT-CAVIAR.pdf

5/18/2005

Nagy IEEE EIT May 2005

19

CAVIAR-FLOWER (continued)



5/18/2005

Nagy IEEE EIT May 2005

20

THE MODEL NEED NOT BE PERFECT!



5/18/2005

Nagy IEEE EIT May 2005

21

AUTOMATIC MODEL CONSTRUCTION



5/18/2005

Nagy IEEE EIT May 2005

22

EASILY CONFUSED FLOWERS



Bellis Perennis
Lawn Daisy, English Daisy



Leucanthemum Vulgare
Ox-eye Daisy



Anemone Canadensis
Windflower,
Canada Anemone



Viola Canadensis
Canada Violet

5/18/2005

Nagy IEEE EIT May 2005

23

CAVIAR Flower Experiments

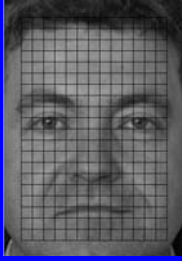
- 30 naïve subjects
- 612 flower pictures of 102 species
- Every interactive mouse click and every automated step recorded in LOG files for detailed analysis

5/18/2005

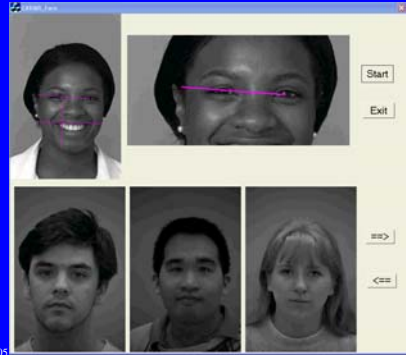
Nagy IEEE EIT May 2005

24

CAVIAR-FACE

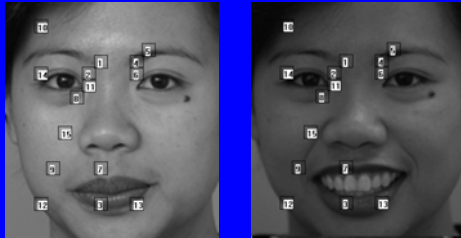


GUI (continued)



FEATURE TEMPLATES (best 15 of 240 candidates)

Most discriminating features near, but not on, eyes.
Single best feature yields 40% accuracy on 200 classes!



CONCLUSIONS

Interactive recognition is twice as **fast** as unaided human, and twice as **accurate** as unaided machine (without years of R&D).

Parsimonious interaction *throughout* the process is better than only at the beginning or end.

Interactive systems can be initialized with a *single training sample per class*, and *improve with use*.

Interaction with images requires a *visible model* that is accessible to both man and machine.

Let both do what they do best.

Leave the human in charge.

STRESS FLOWER DATABASE

- 320 by 240 pixel pictures
- Highly variable illumination, and complex background
- 216 samples from 29 classes for development
- 612 samples from 102 classes for evaluation
- Most (digital) photos from New England Wildflower Garden

CAVIAR-FACE INTERACTIONS (6 subjects, 200 faces)

SELECT	RANK ORDER	
	ADJUST	BROWSE
50.3% (2.3 sec)	48.7%	1.0% (7.7 sec)
19.7% (7.7 sec)	28.3%	0.7% (16.1 sec)
13.3% (10.6 sec)	15.0%	0.0%
4.7% (14.4 sec)	9.3%	1.0% (42.6 sec)
5.3% (16.6 sec)	3.7%	0.3% (23.2 sec)
2.0% (19.6 sec)	1.0%	0.7% (33.2 sec)
0.3% (42.0 sec)	0.7%	0.0%
0.3% (34.7 sec)	0.3%	0.0%
0.0%	0.0%	0.3% (49.8 sec)

M-CAVIAR Classification Example



(1) Automatic is unsuccessful as the flower is out of focus. (2) petal number changed to 5 & the re-estimated rank order and rose-curve instance is displayed. (3) the inner radius and phase changed to fit the flower better and the correct candidate appears.

5/18/2005

Nagy IEEE EIT May 2005

31

CAVIAR-FLOWER



5/18/2005

Nagy IEEE EIT May 2005

32

COMPUTER BASED INTERACTIVE RETRIEVAL vs. CAVIAR

CBIR	CAVIAR
Subjective retrieval	Objective classification
User judges retrieval results	Statistical decision boundary
User weights features	Machine weights features
Broad domain	Narrow domain
Relevance feedback	Relevance feedback
	Model adjustment

5/18/2005

Nagy IEEE EIT May 2005

33

Flower Database (1)



5/18/2005

34

Flower Database (2)



5/18/2005

Nagy IEEE EIT May 2005

35

Flower Database (3)



5/18/2005

Nagy IEEE EIT May 2005

36

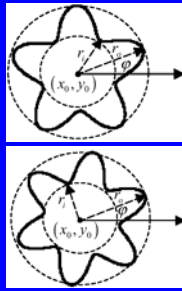
ROSE CURVE MODEL

- Parametric curve with six parameters.

$$\rho = \frac{r_o + r_i}{2} + \frac{r_o - r_i}{2} \cos(n\theta + n\varphi)$$

$$= a + b \cos(n\theta + n\varphi)$$

- Flowers are composed of petals, which have circular symmetry.
- When $n=0$, rose curve reduces to circle.



5/18/2005

Nagy IEEE EIT May 2005

37

Examples of visual pattern recognition

- Bar codes (e.g., UPC) ✓
- Normal printed matter ✓
- Motivated hand print ✓
- Fingerprints ✓
- Gross thematic maps from satellite pics ✓
- Industrial part and assembly inspection ?
- Target recognition ?
- Printed matter in complex formats ?
- Degraded (faxed, copied) printed matter ?
- Sloppy or archaic handwriting
- Detailed thematic maps
- Micrographs, X-rays, skin lesions
- Faces (*lighting, pose, expression, aging*)
- Cryptic cats, birds, fish, flowers, ...

5/18/2005

Nagy IEEE EIT May 2005

38

CAVIAR Experimental Protocol

Experiment Type	# of Subjects	Training Samples	Test Sample	Notes
I	6	1,2,3,4,5	6	Browsing-only with 5 reference samples
II	6	1,2,3,4,5	6	Interactive with 5 training samples
III	6	1	2,3	Interactive with 1 training sample
IV	6	1,2*,3*	4,5	Interactive with 1 training sample + results of III
V	6	1,2*,3*,4*5*	6	Interactive with 1 training sample + results of III, IV

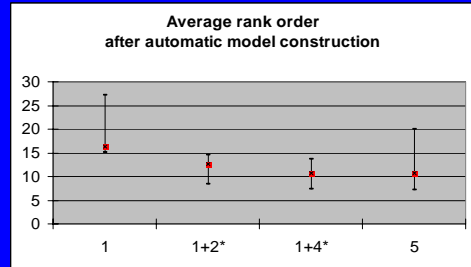
* samples initially without labels

5/18/2005

Nagy IEEE EIT May 2005

39

SYSTEM ADAPTATION

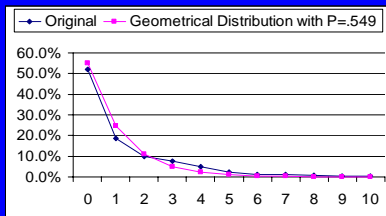


5/18/2005

Nagy IEEE EIT May 2005

40

Finite State Machine model of interaction



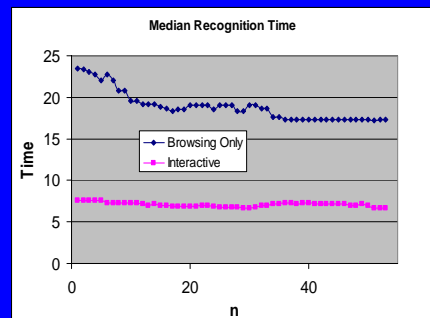
- 52% samples are immediately confirmed.
- 90% samples are identified after 3 adjustments.
- The probability of success on each adjustment is ~0.5.

5/18/2005

Nagy IEEE EIT May 2005

41

HUMAN LEARNING



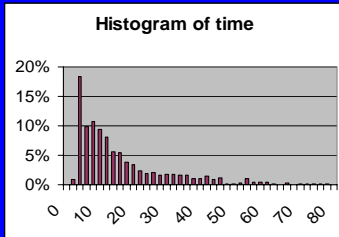
5/18/2005

Nagy IEEE EIT May 2005

42

ENROLLMENT:
REFERENCE DATA SEGMENTED
WITH INTERACTIVE CORRECTION

- 15.2 seconds per picture (5.7 seed pixels),
- 1078 flowers from 113 species



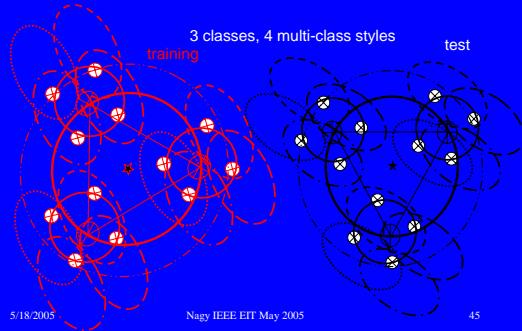
Welcome to
**Computer Assisted Visual
InterActive Recognition
(CAVIAR)**

CAVIAR is an interactive flower classification program. By interacting with the computer, we hope that you can recognize flowers more accurately than a computer can by itself, and faster than you can without computer help.

RPI ECSE DocLab
Jie Zou, Borjan Gagoski, George Nagy

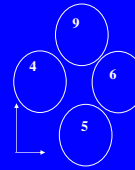


WEAKLY CONSTRAINED DATA
given $p(x)$, find $p(y)$, where $y=g(x)$



Are weak constraints enough?

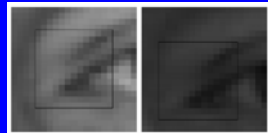
Training



Test



Search over a
5x5 window



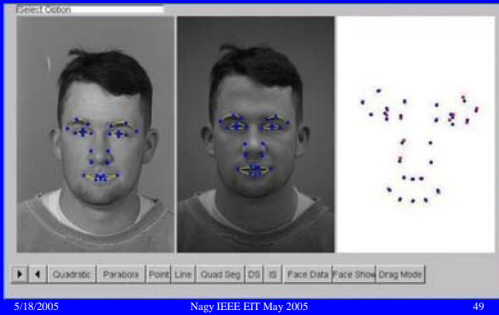
Probe

G1 G4
Gallery

T E M P L A T E S	Gallery (reference) faces									
	G1		G2		G3		G4		G5	
	Similarity	Rank	Similarity	Rank	Similarity	Rank	Similarity	Rank	Similarity	Rank
P1	0.999501	1	0.997885	5	0.997886	4	0.998195	2	0.998056	3
P2	0.997412	2	0.997273	3	0.997989	1	0.996801	5	0.997120	4
P3	0.970771	2	0.960403	5	0.964492	4	0.975555	1	0.970332	3
Borda Count	5		13		9		8		10	
Final Rank	1		5		3		2		4	

CAVIAR-FACE: FIDUCIAL POINTS AFTER SIMILARITY TRANSFORM

Matt Green



5/18/2005

Nagy IEEE EIT May 2005

49

MISRECOGNIZED FACES

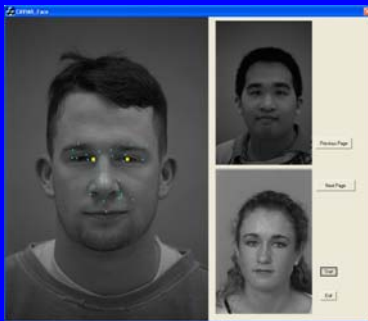


5/18/2005

Nagy IEEE EIT May 2005

50

CAVIAR-FACE (BAD PUPIL LOCATION)

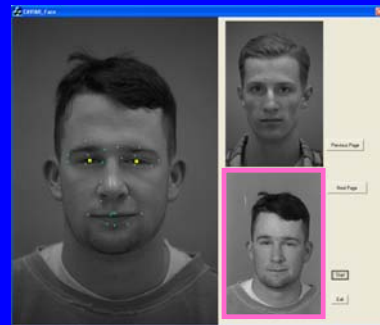


5/18/2005

Nagy IEEE EIT May 2005

51

CAVIAR-FACE (GOOD PUPIL LOCATION)



5/18/2005

Nagy IEEE EIT May 2005

52

FEATURE EXTRACTION AND CLASSIFICATION

Affine **size normalization** based on model
 Local **histogram equalization** on template surround
 Cosine **similarity measure** on 11x11 feature templates
 5x5 **search window** for each template
Features selected by agglomerative search
Borda Count classifier based on rank order
 (usually only five features required for Top-3)
 Difficult face-pairs require more features,
 but only extracted from leading candidates
 Other experiments on **pose, expression, aging, ...**

5/18/2005

Nagy IEEE EIT May 2005

53

DECISION-DIRECTED ADAPTATION

RESULTS:

<u>Year</u>	<u>Collaborator</u>	<u>Data</u>	<u>#_classes</u>	<u>d</u>	<u>Gain</u>
1966	Shelton	12-font typescript	26	96	5.0X
1994	Baird	100-font print	96	512	2.5X
2002	Harsha V.	NIST hand-print	10	50	1.8X
2003	El-Nasan	curved handwriting	100	42	4.0X
2004	Zou	flowers	102	8	1.2X

5/18/2005

Nagy IEEE EIT May 2005

54

PROPERTIES OF NATURAL PATTERNS

- Photographed: varied pose, "expression", *lighting*
- Evolved without regard for discriminability
- Arbitrary order (no linguistic context)
- The number of classes is subjective
- Hierarchical taxonomy based on non-visual aspects
- Ancillary non-visual information is often available