

How Many Words is a Picture Worth?

Pictures, Pixels and Words

Professor Ralph Jacobson PhD., CChem FRSC., ASIS Hon FRPS
Computational Vision and Imaging Technology (CVIT)
School of Computer Science & Engineering , University of Westminster

**Imaging Science Group RPS Symposium:
Good Picture - 2022 ,
University of Westminster London, December 2022**

Content

- ❖ Introduction
- ❖ Background and Context
- ❖ Preliminary work and approaches
- ❖ Results of Preliminary work
- ❖ Extension to image attributes
- ❖ (Future work on Image Cells)
- ❖ Discussion
- ❖ Conclusions

Introduction

Question prompted by Gary Evan's Presentation at GP2019.

Any number of words not to be taken too literally!

Observations carried out during the Covid 19 Lockdowns.

Quest for answer(s) based on:

Image Science and related disciplines.

A case and pilot study of work in progress.

Results are for the images and methodology used.

One Picture is worth 10,000 words

PRINTERS' INK Mar. 10, 1927

PRINTERS' INK Mar. 10, 1927

達萬言 畫意能

CHINESE PROVERB
One picture is worth
ten thousand words

ROYAL
BAKING
POWDER
ABSOLUTELY PURE

ROYAL
BAKING POWDER

Avoid Substitutes

NO ALUM

"Make a Cake for Bobby"

—that's what this car card said *every day* to many millions of women. It reminded all mothers *every day* of a sure way to give a treat to their own children. And hundreds of thousands got an extra thrill with their next cake making because of the happy expression of the boy on the car card.

The moral of this story is that the same influence could not be created even with the same picture in any other advertising medium.

In the magazines, the reminder would not be often enough to change the average housewife's baking habit. In the newspapers, with no color, there would be no appetite appeal. On a twenty-four sheet poster, seen for only a few seconds at a time, the great appeal of the expression on the boy's face would be lost.

H. Barnard
National Advertising Manager.

STREET RAILWAYS ADVERTISING CO.

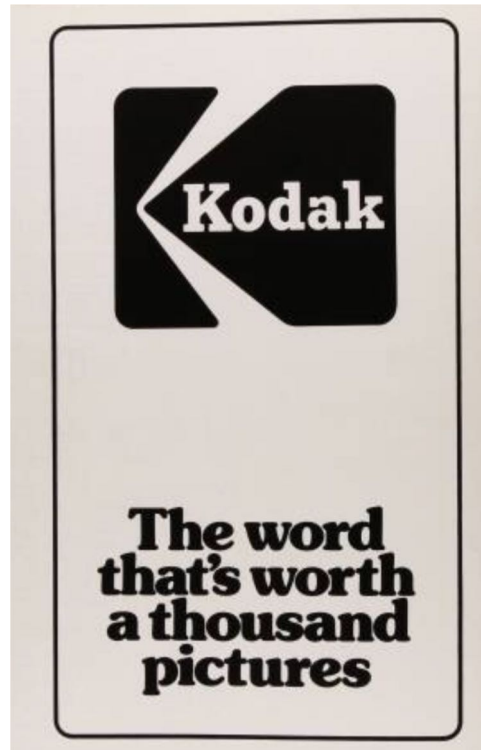
Is it 1000 words?

In March 1911, the Syracuse Advertising Men's Club held a banquet to discuss journalism and publicity. In an article covering this event, the author quoted Tess Flanders as saying:

"Use a picture. It's worth a thousand words."



Advertising: Play on Words



Music

Song by: Bread, Telly Salvalas, Shirley Bassey etc. :

If a picture paints a thousand words

Then why can't I paint you?

The words will never show the you I've come to know

If a face could launch a thousand ships

Then where am I to go?

Songwriter: David Gates. 1971

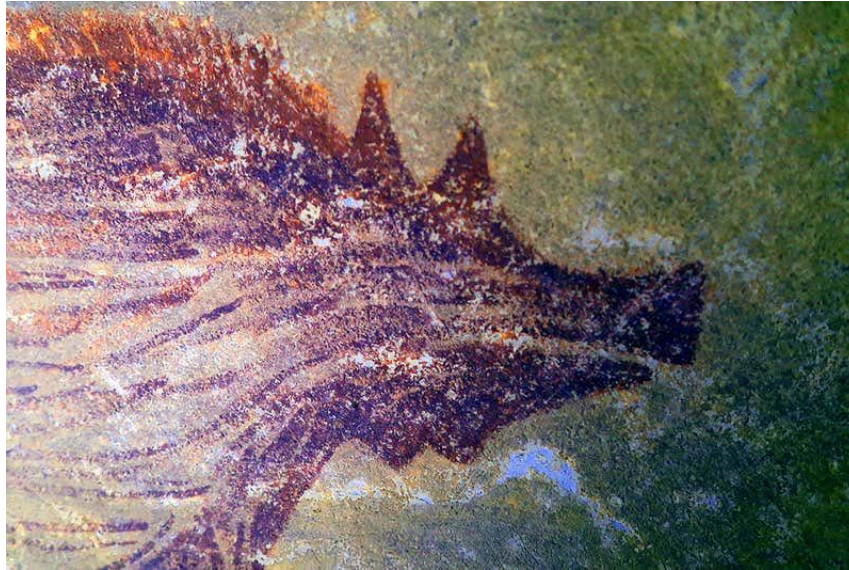
News Item



Boris Johnson raising a glass at a leaving party on 13th November 2020

Photos like this tell a **thousand** words

The very Beginning; No Words!



This painting of a wild pig in the Leang Tedongnge cave on the Indonesian island of Sulawesi is thought to be the oldest representational art in the world (at least 45,000 years ago)

A Few Previous Investigations

Is a picture worth a thousand words? (Bishop 1977)

A picture is not always worth a thousand words. (Willows 1978)

Why a diagram is (sometimes) worth ten thousand words.
(Larkin & Simon 1987)

When a picture is not worth a thousand words. (DeLoache & Marzolf 1992)

A picture is worth a thousand words, but that's the problem.
(Gleitman & Gleitman 1992)

When is an illustration worth ten thousand words? (Mayer & Gallini 1990)

Correction: A Picture is Worth 84.1 Words. (Blackwell, 1997)

An image is worth more than 100,000 words, rather than 1,000 words!
(Kulkarni 2000)

Iconic Images: No Words!

“.....Pictures may well be worth a thousand words but sometimes they are so powerful there are no words. These photos may be shocking, beautiful or both but they speak a different language because they transcend verbal communication..... “

John Dabell, Picture News 31 October 2017

Picture and Numbers



“Claude Shannon has one equation that boils down the whole essence of communication and to me it feels so perfect to look at something in such a clear way”

The RPS Journal, Vol 160, 280 (2020)

Entropy proposed by Shannon in 1948 which he originally termed ***uncertainty***.

Methodology

Use of thresholds

Selection of Images

Image Pipeline

Panel of Observers, Questions asked

Choice of metrics

Results

Discussion and Conclusions

(Including diversions and future approaches)

Thresholds

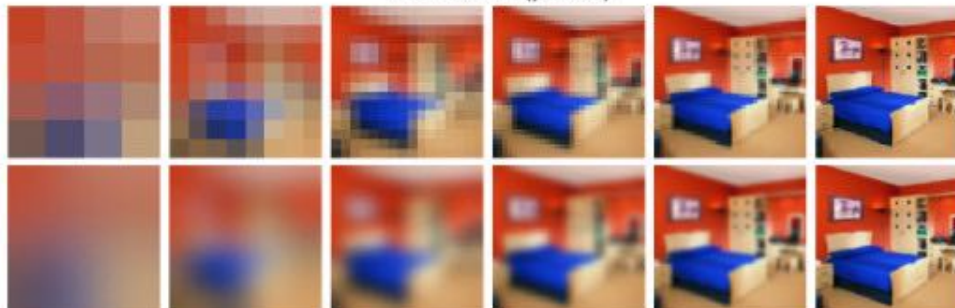
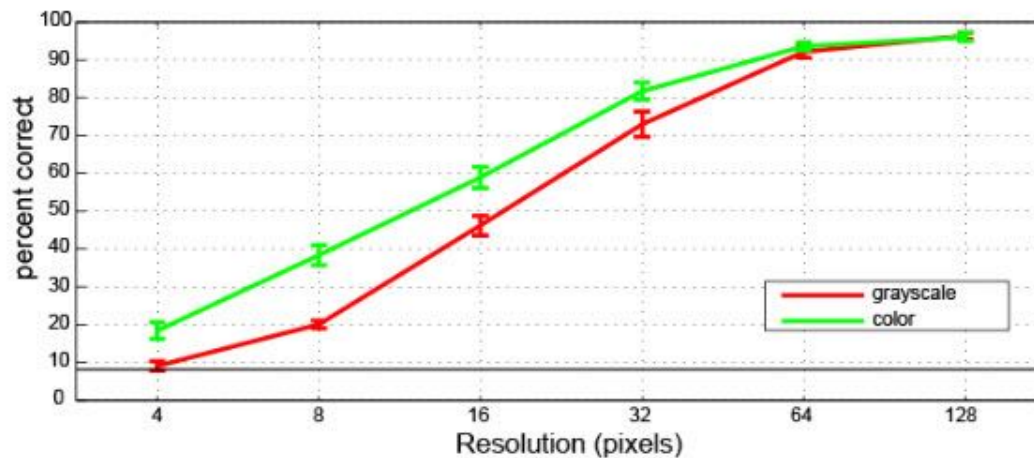
The answer to the question has no upper limit.

Modern Imaging applications need less data to speed computation
(e.g. autonomous vehicles, target recognition).

Many applications have to use low resolution and low quality images
(e.g. surveillance, face recognition).

Threshold Criteria

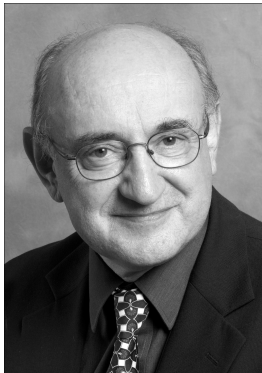
Minimum No of Pixels for Scene Categorization



32 x 32 (1024 pixels)

Selection of Images

1



3



2



4



5



6



7



8

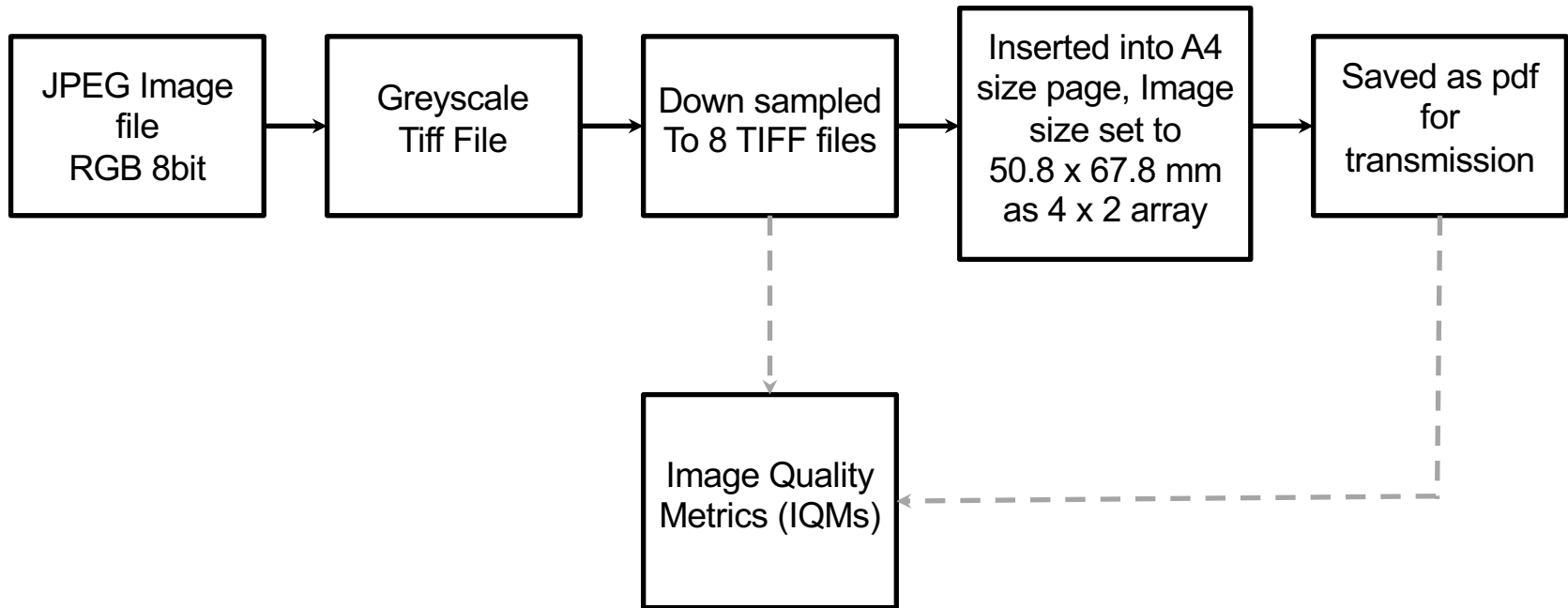


9

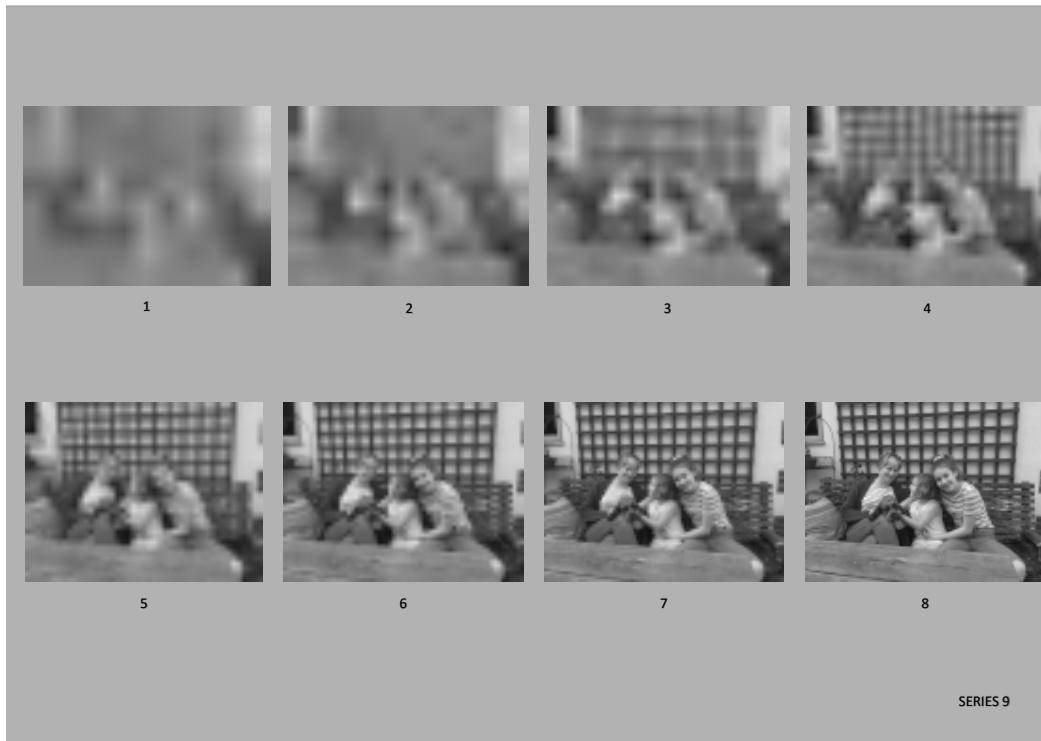


10

Image Pipeline



Down Sampled Images Sent to Observers



Observers and Viewing conditions

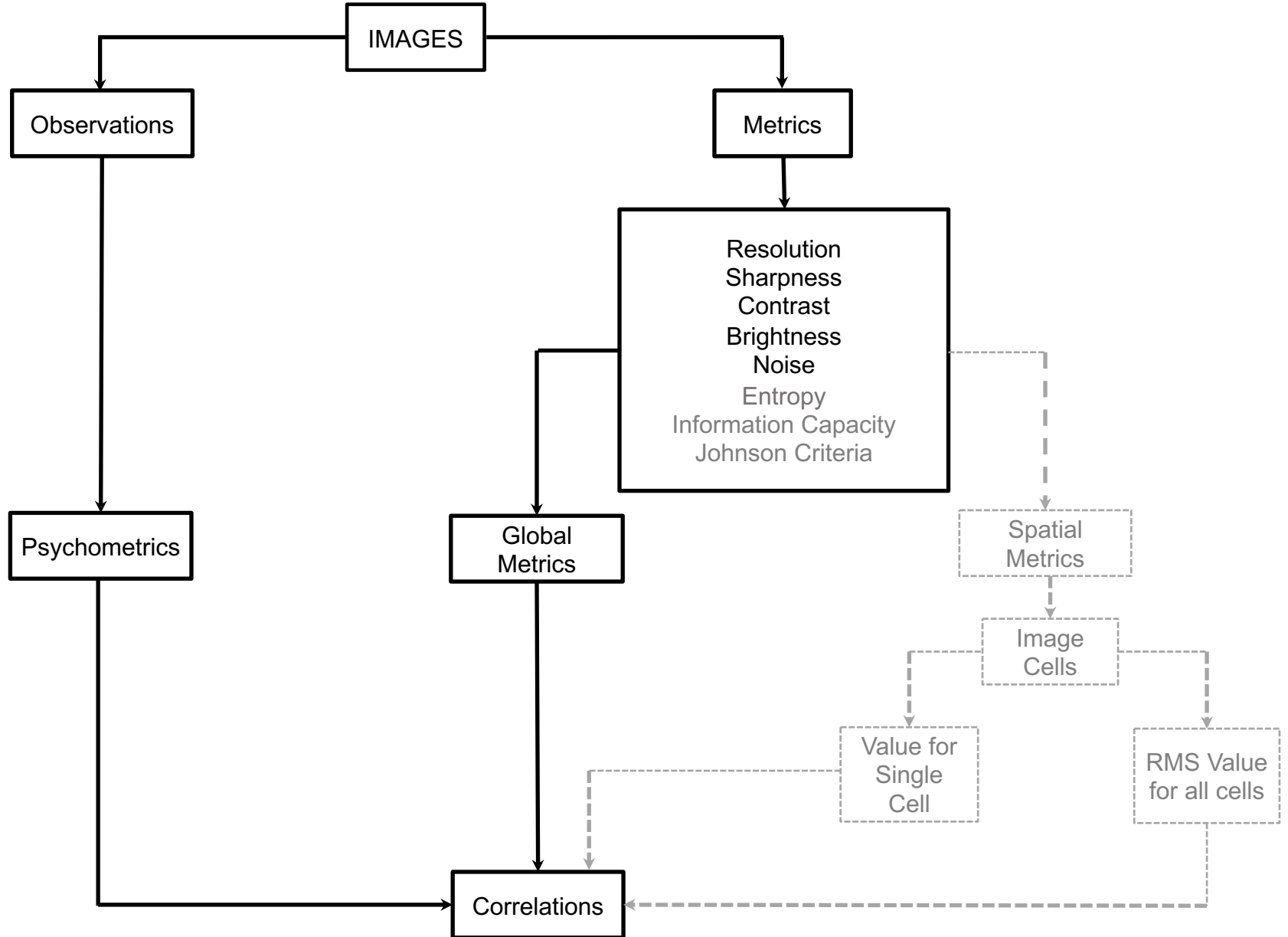
The observers were asked to view the images in full display mode at their normal viewing distance and not to use zooming.

Two sets of images were e-mailed to observers over a period of around 10 weeks.

Observer Tasks

1. Which picture shows just sufficient information for you to decide what the picture represents (e.g. a portrait, a landscape, a car, a dog etc.)?
2. Which picture shows just sufficient information for you to identify the detail of its content (e.g. to identify a person in the portrait, the content of the landscape, the type of dog, flower, car etc.) ?
3. What screen size are you using, measured as the diagonal in mm?
4. When the observers had viewed all images they also were asked to name any particular images that they found easy or difficult to judge.

METRICS FLOW CHART



Software used

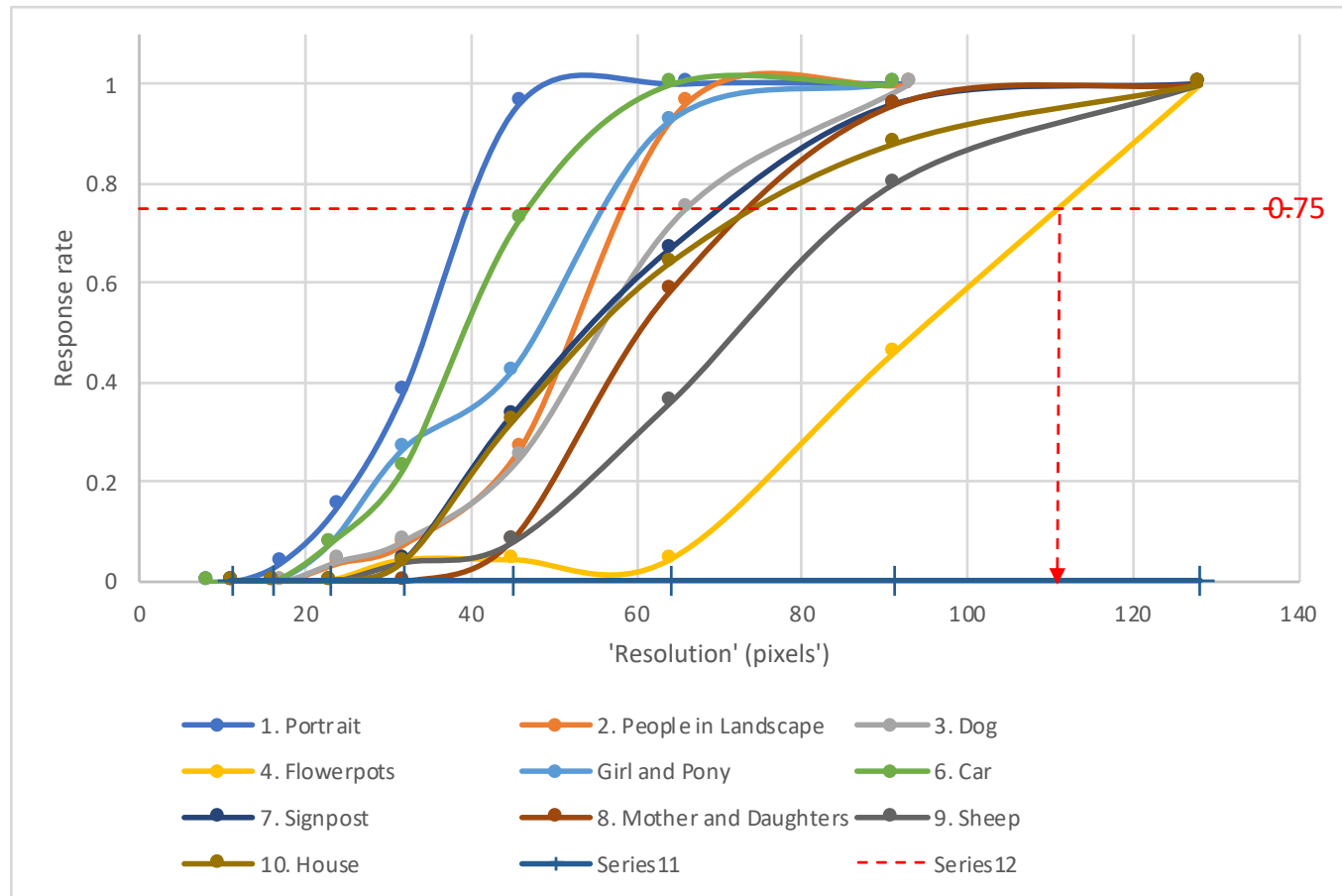
- *GIMP 2.10.14 GNU* Image Manipulation Program for file conversions and down sampling.
- *Microsoft Excel V 16.16.27* for statistical tests, supplemented by the add-in *Real Statistics Resource Pack for Mac 2011 V 5.11*.
- *Microsoft Excel, GNU Octave 6.2.0* and *Matlab R2002b* for computation and calculations.
- All data were subjected to the Shapiro-Wilk test for normality to decide on whether to use parametric or non-parametric correlations. (Shapiro and Wilk,1965).

Calculation of results

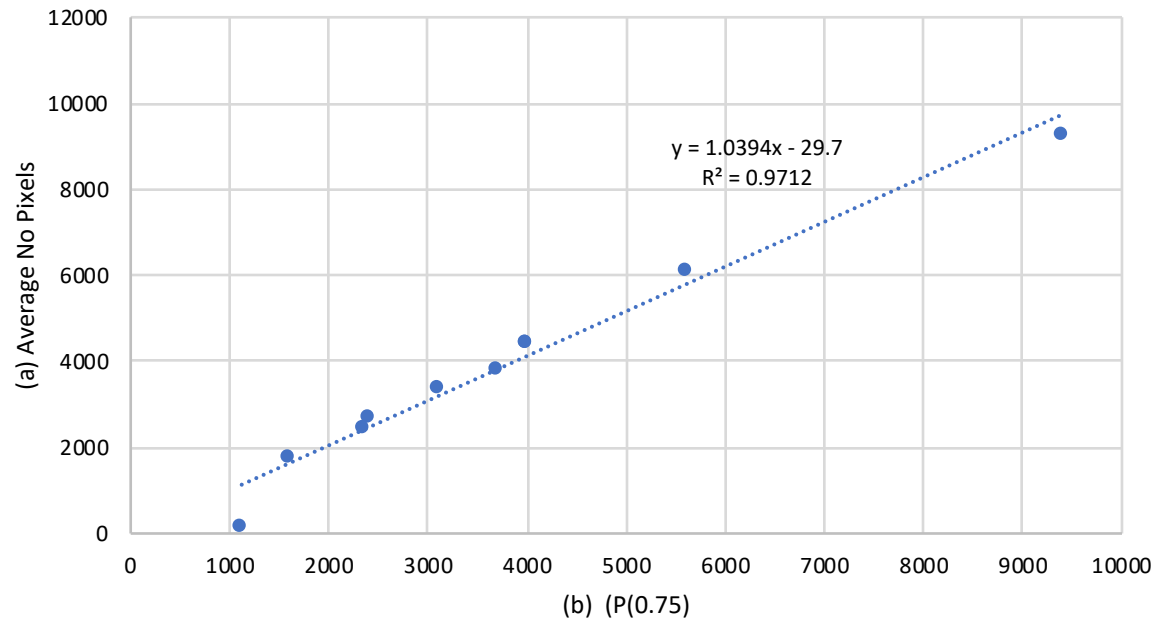
For these forced choice observations:

- a) The average pixel resolutions (numbers of pixels) for the images selected by the observers.
- b) From curves of response rate (0 to 1) vs pixel resolution at 75% level ($P = 0.75$).

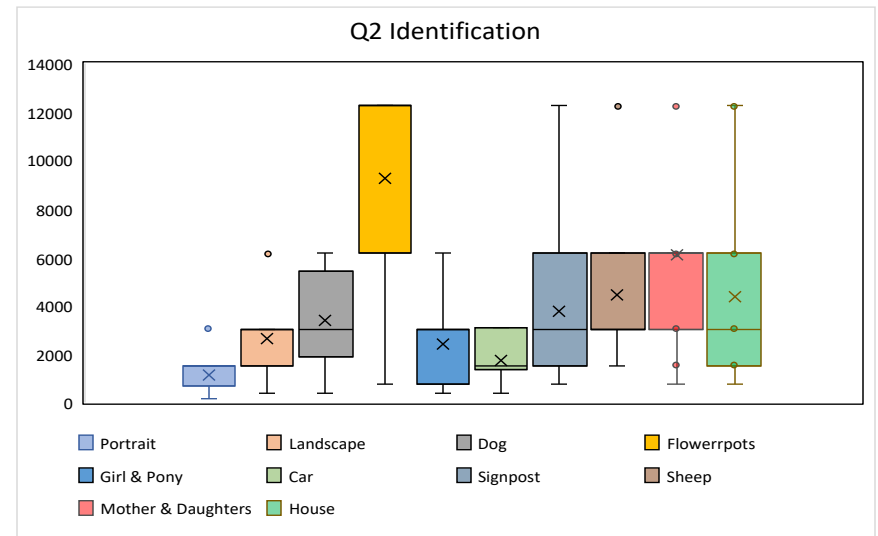
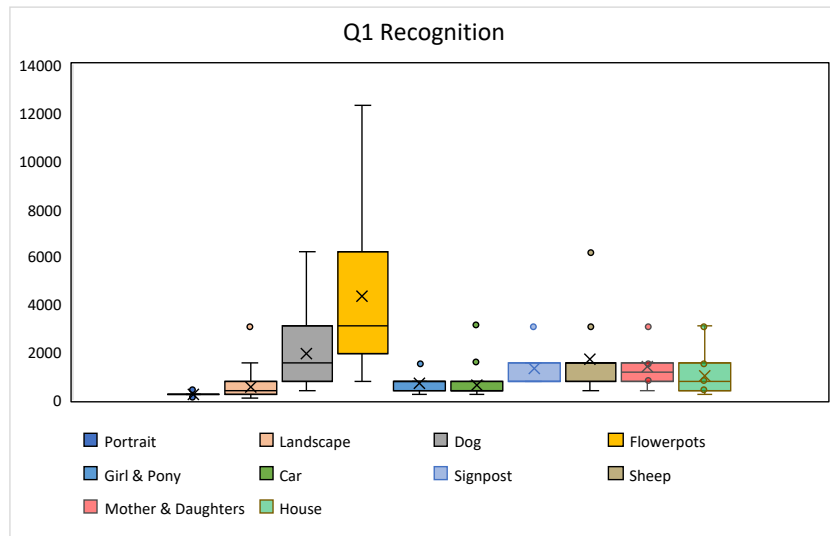
Response Rates and Resolution



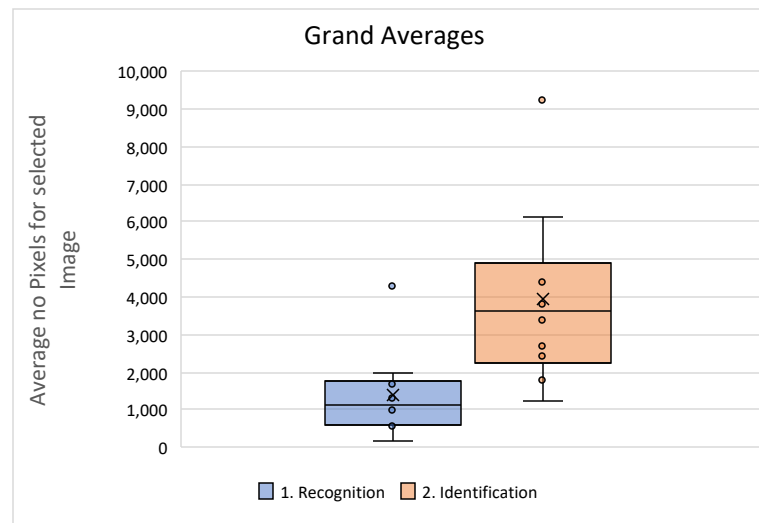
Response rates and Averages



Results for All Observers and Images



All Results All Observers



Conversion between Word and Image Data?

- Smallest unit of image data is a pixel
- Smallest unit of word data is a character
- One Character = 1 byte (8bits)
- One pixel typically consists of 8 bits (1 byte) for a Black and White image.

Hence one pixel is equivalent to one character.

How do we convert Pixels to words? (need to know average word length)

How Many Characters in a Word?

Word Lengths:

“And here is the breakdown of mentions (in millions) by word length (looking like a Poisson distribution). The average is 4.79 letters per word, and 80% are between 2 and 7 letters long”.

*English Letter Frequency Counts: Mayzner Revisited or **ETAOIN SRHLDCU** (Norvig, 2012)*

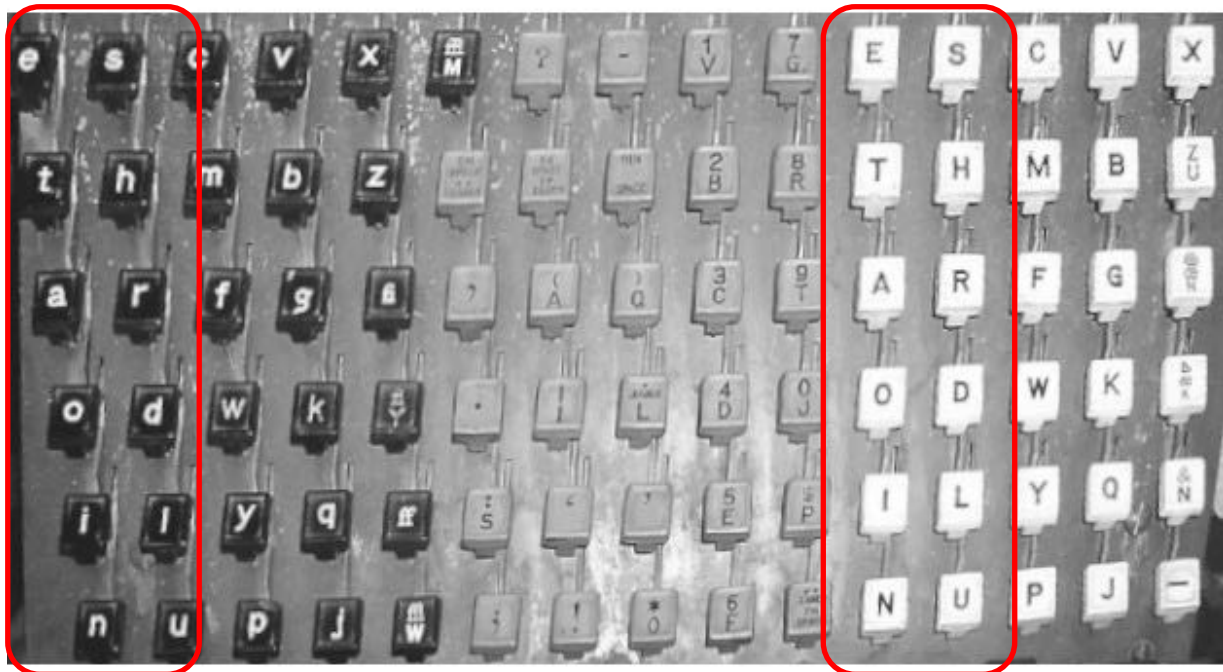
Who or What is **ETAOIN SRHLDCU** ?

A Relative of QWERTY:

The Linotype keyboard.

For the production of lines of relief printing type in
letterpress printing.

In use for around 100 years until the mid 1970s'



Pixels to words

one letter = one pixel,

one word = 4.79 letters = 4.79 pixels,

one pixel = 1 / 4.79 words

How Many Words is a Picture Worth?

Scenes and numbers of pixels	No Words
All scenes: recognition/identification, 1,371/3,972	286/ 819
Portrait: recognition/identification, 185/1,220	39/ 322
All scenes (Torralba) : recognition/identification, 32 x 32	214/
Police trained 'super-recognizers' : face recognition, 30 x 45	/ 282
Absolute range: min for recognition/max for identification, 204/12,288	43/2565
Colour image (Kulkarni): 'high quality', 512 x 512	100,000

All scenes no of Words = 819

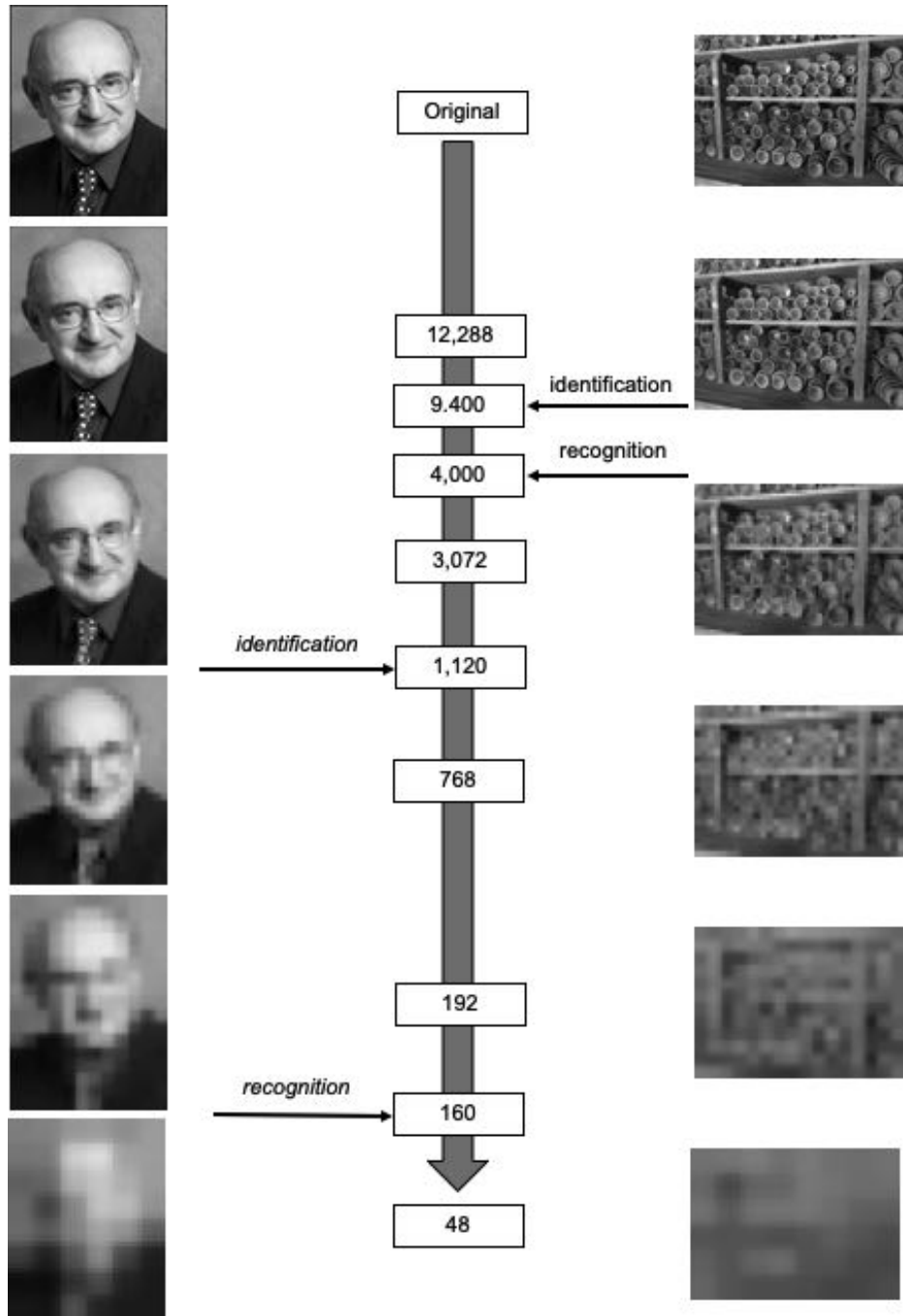
Average no of words from absolute minimum and maximum = 1,303

Average of these two = 1,061 !!!!

All Closer to 1,000 than 10,000WOW!

Pixel Resolution

Scene Dependency Causes?



Different results for different scenes?

1) Cognitive issues

2) Image attributes

Intuitively:

1) Resolution

2) Sharpness

3) Contrast

4) Brightness

5) Noise

3) Others (for applications and benchmarking?)

Entropy, Information Capacity, Johnson

Criterion for 2) & 3) this work:

Must be accessible, easy to calculate and widely used

Metrics used (IQM)

Resolution (Pixels, Cycles/mm)

Sharpness (RMS edge measure)

Contrast (RMS contrast)

Brightness (Mean luminance)

Noise (Signal-to-noise ratio)

Entropy (histogram (PDF))

Information Capacity (pixels&values)

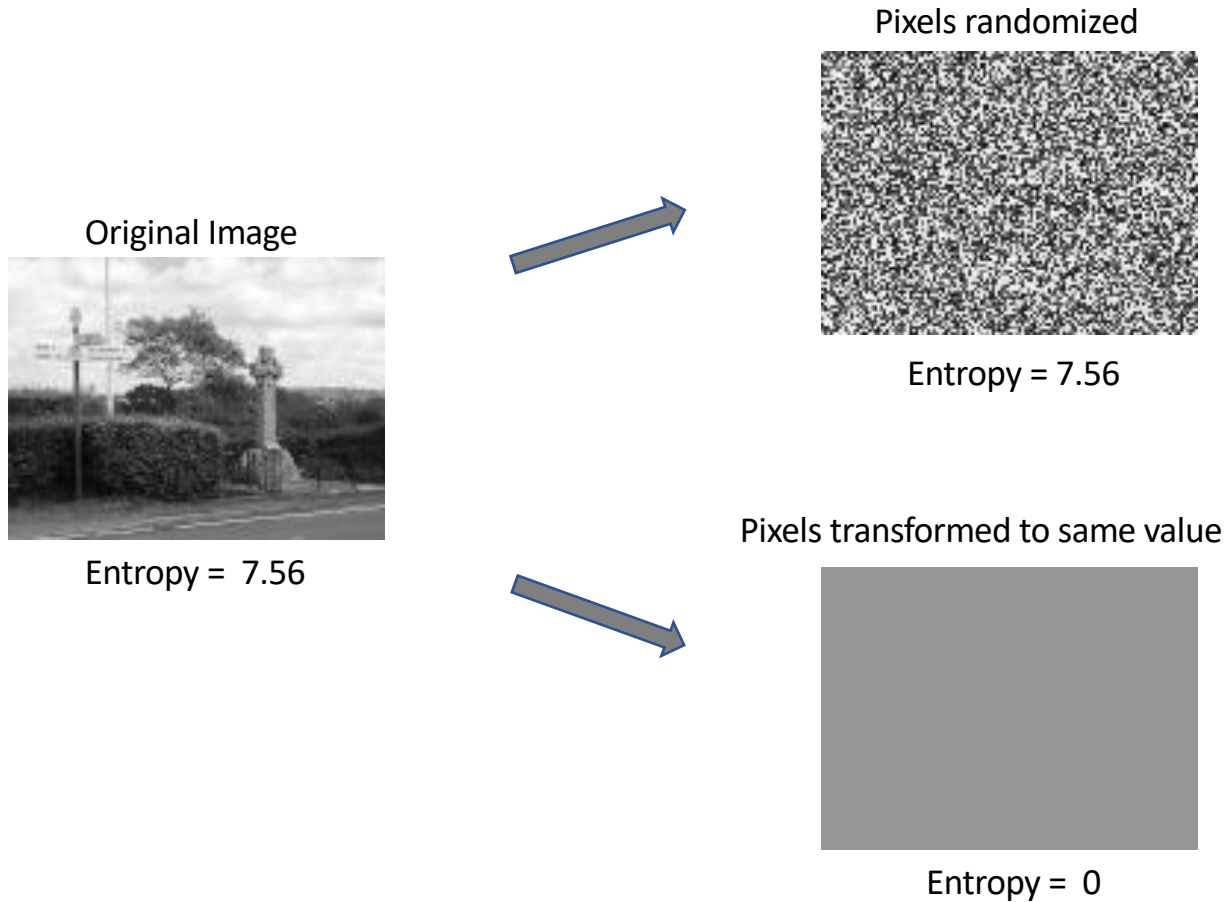
SCR (Signal-to-clutter ratio)

Johnson Criteria (pixels on 'target')

RMS: root-mean-square

PDF: Probability Density Function

Entropy



Higher Entropy: Less structured, more 'randomness', more information

Lower Entropy: highly structured, little 'randomness' little information

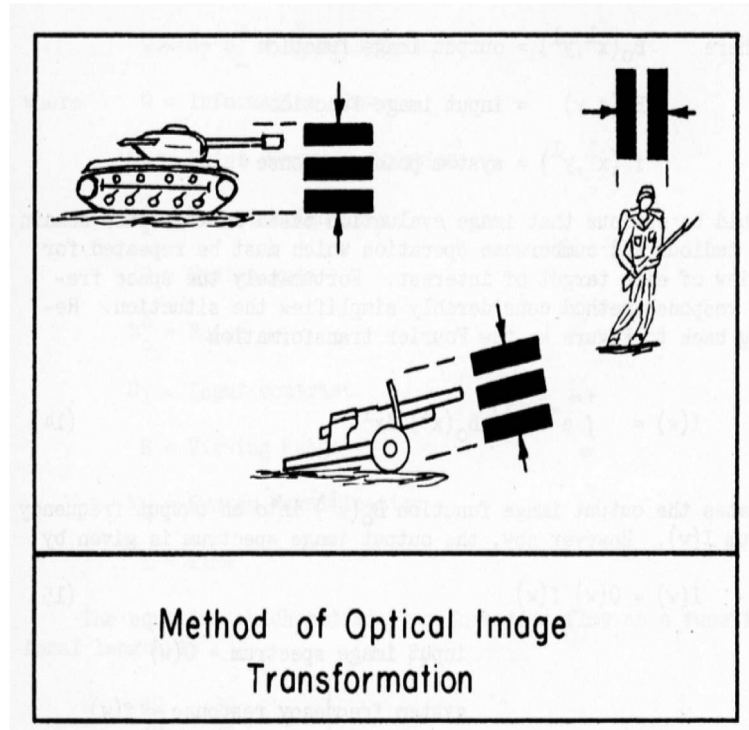
Reversibility: Process with no change in entropy is easily reversible ?

DOES NOT INVOLVE SPATIAL (POSITIONAL) INFORMATION

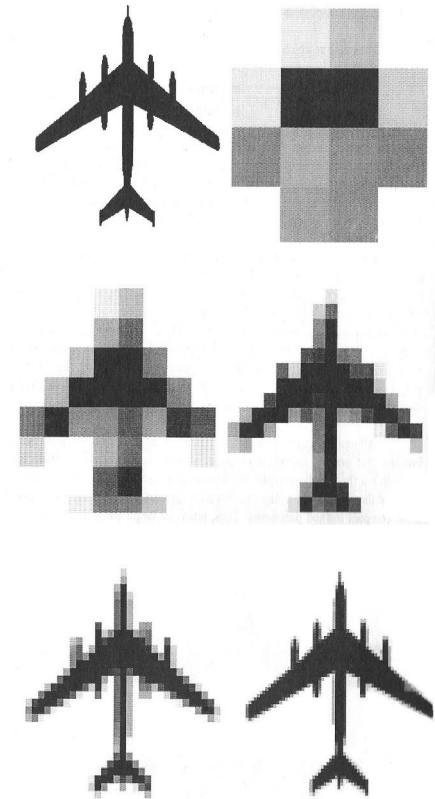
“In conclusion to this section it seems that von Neumann's comment “...*no one knows what entropy really is*” accurately describes the current situation with respect to Shannon's information theory applied to signals, and images in particular”.

K. G. Larkin, 2016

Johnson Criteria



Johnson, 1958



Bear Bomber, Holst, 2002

Scene Dependency

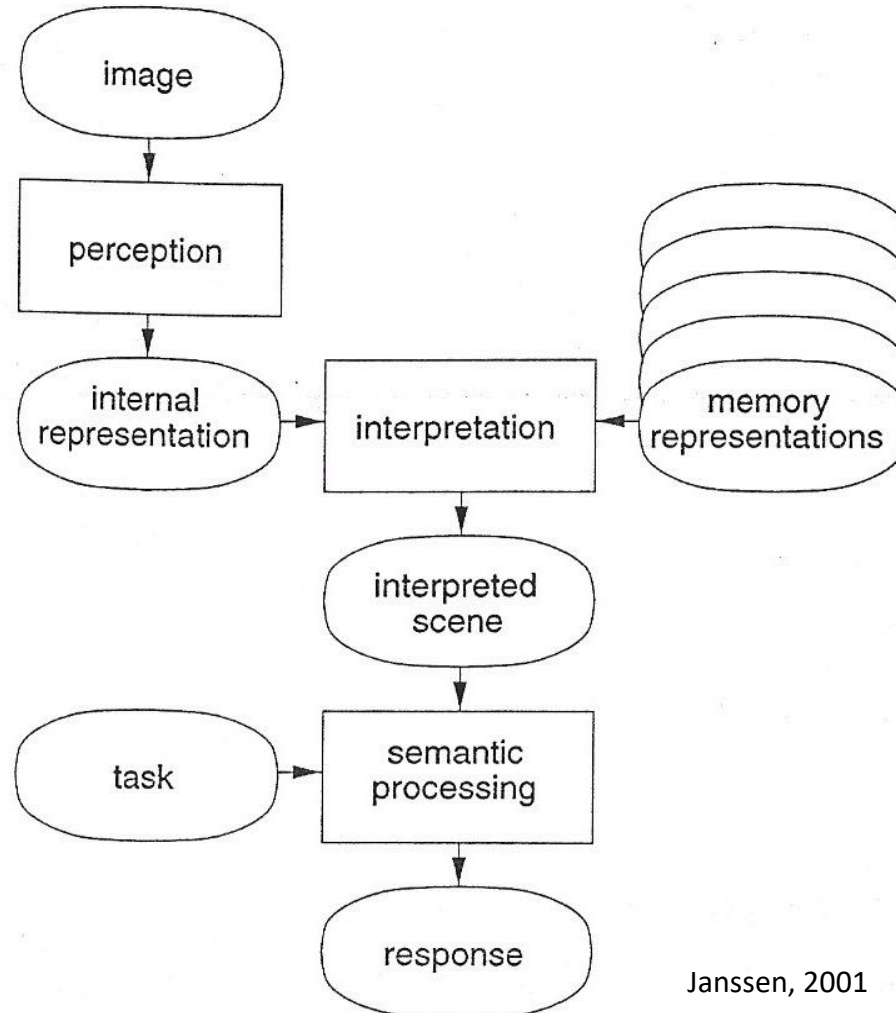
Observer preferences

Visuo-cognitive effects

Visibility of artefact or attribute

Effect of image processing algorithms
(scene susceptibility)

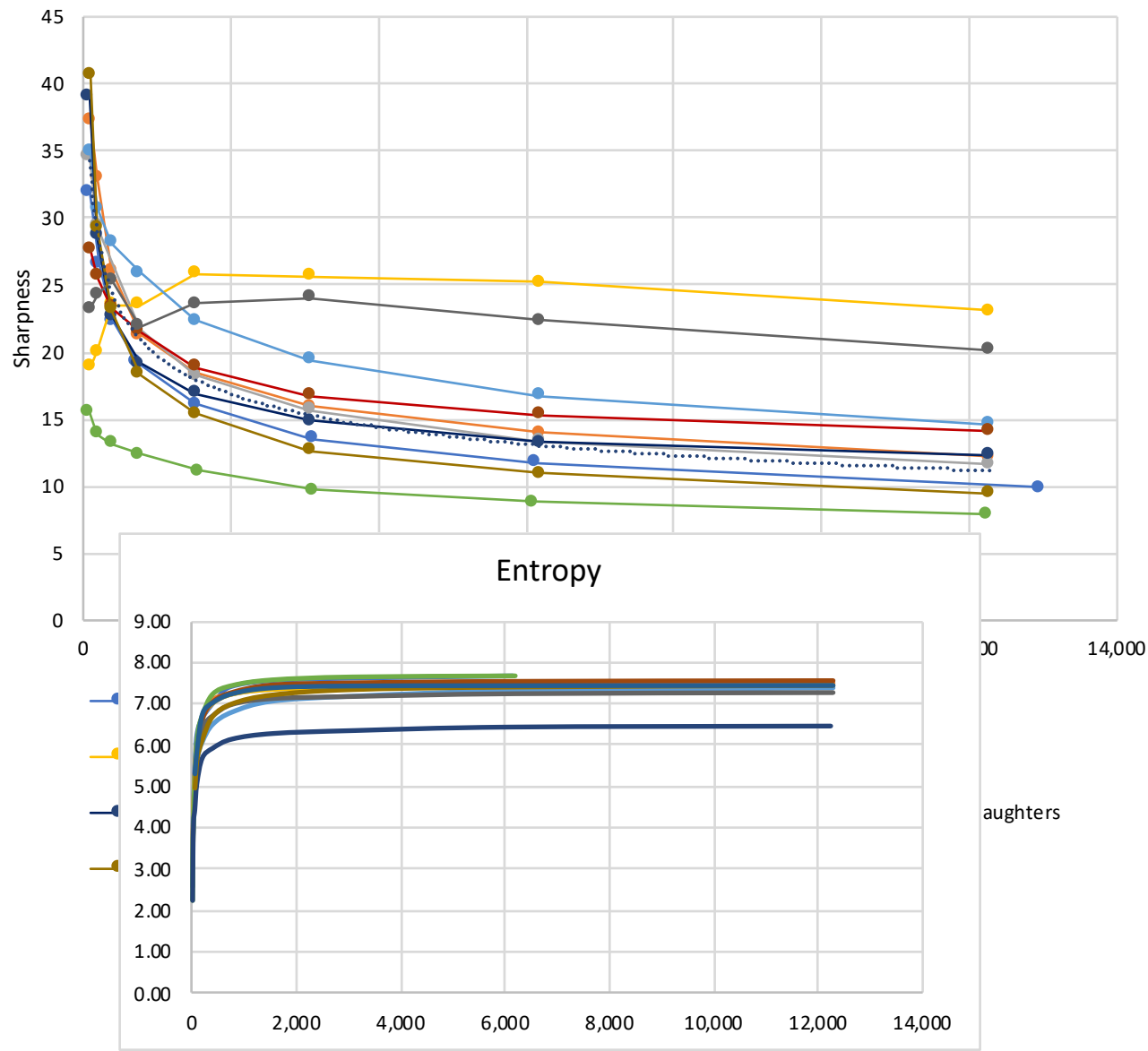
Visuo Cognitive Effects



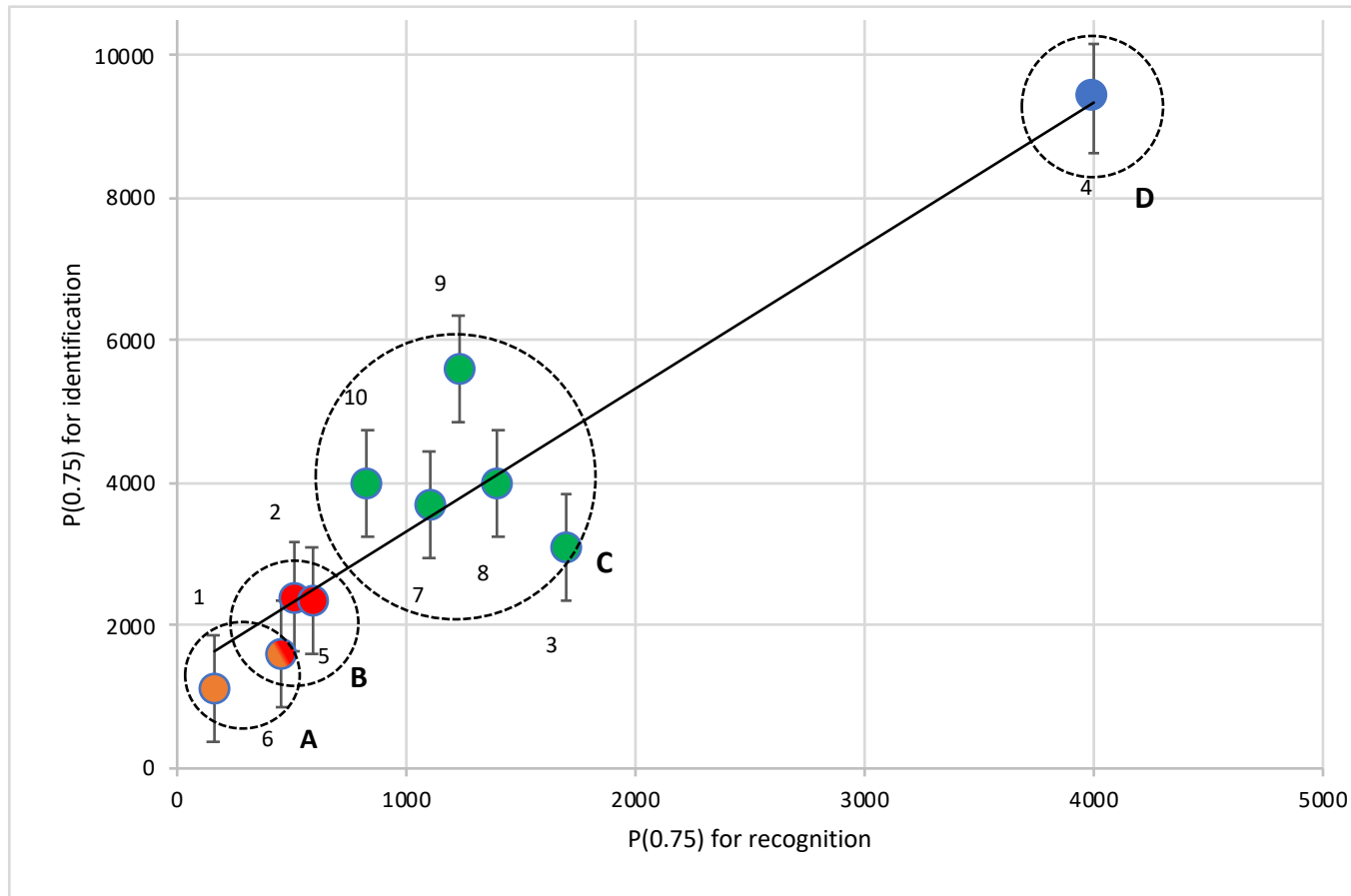
Visibility of artefact or attribute



Scene Susceptibility (Sharpness)



Clusters



Pearson correlation coefficient ($R^2 = 0.849$)

Spearman's correlation coefficient ($\rho(8) = 0.83$, $p = 0.003$).

Observation Results for Clusters (Identification)

Cluster	Image	Resolution P(0.75)		IQR		Difficulty	
		Value	mean	Value	mean	Value	mean
A: very low	1. Portrait	1120	1360	782	975	-5	-2.5
	6. Car	1600		1,167		0	
B: Low	2. Landscape	2400	2375	1,542	1828	3	1.5
	5. Girl & Pony	2350		2,114		0	
C: intermediate	3. Dog	3100	4080	1,165	3343	4	2
	7. Signpost	3700		4,658		1	
	8. Sheep	4000		3,116		1	
	9. Mother & Daughters	4000		3,116		2	
	10. House	5600		4,658		1	
D: very high	4. Flowerpots	9400	9400	6,100	6100	8	8
Mean	All images	3727	4304	2842	3061	2	2

Scenes of Cluster C can be considered to be 'average scenes'

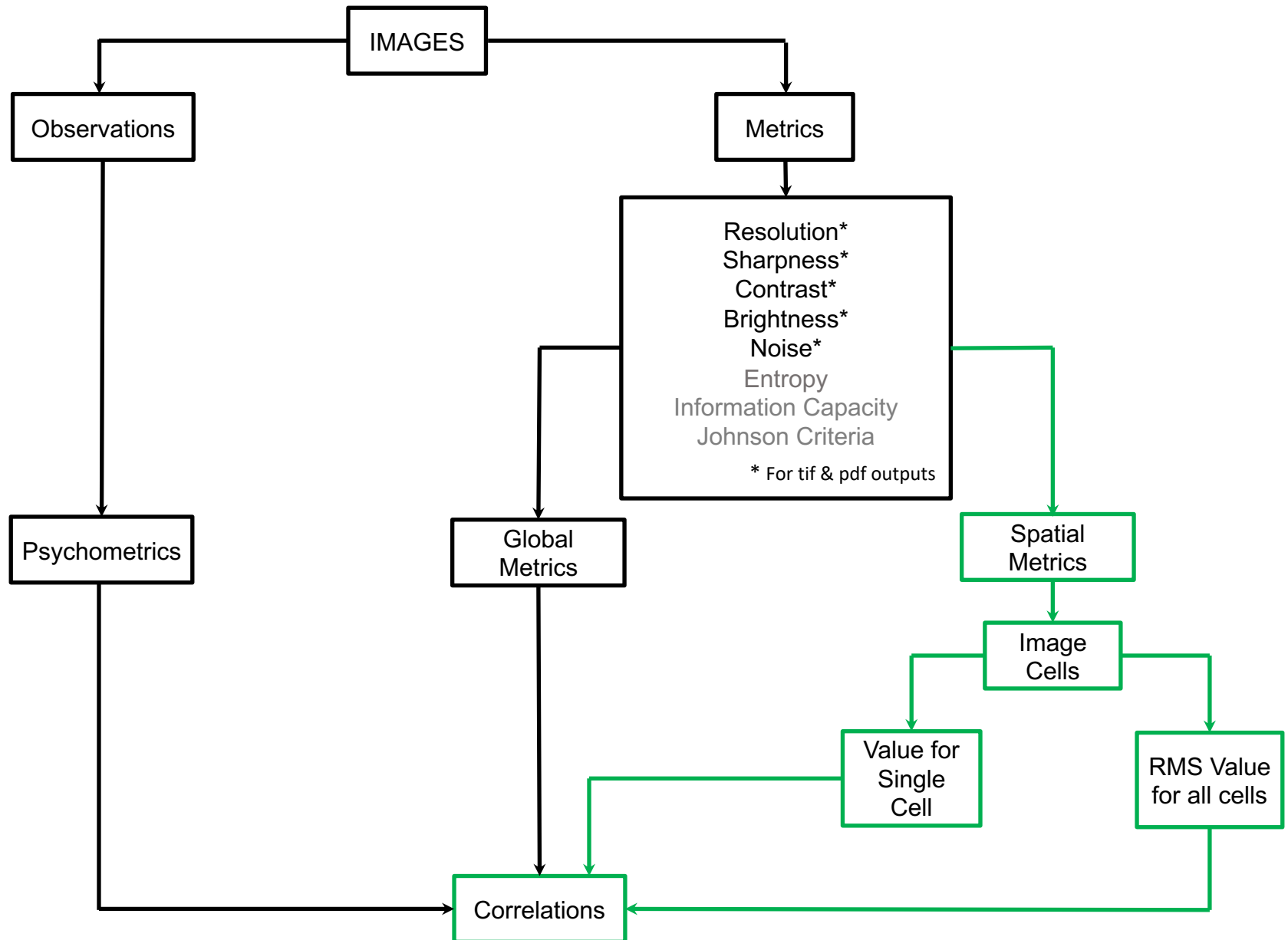
No strong correlations were found for any of the image attributes (other than resolution)

Metrics and Spearman correlations for Global Image Attributes

<i>Metric</i>	<i>Pixel Resolution (P(0.75))</i>	<i>IQR</i>	<i>Difficulty</i>
<i>Resolution (cycles/mm)</i>	1.000	0.826	0.691
<i>Entropy</i>	-0.365	-0.079	-0.332
<i>Contrast</i>	-0.036	0.159	-0.345
<i>Sharpness</i>	0.438	0.256	0.308
<i>SNR</i>	0.432	0.195	0.302
<i>Luminance</i>	-0.122	-0.098	0.074

(95% chance of being significant for values > 0.564)

METRICS FLOW CHART (and future work)



Conclusions for Global Results

- A picture is worth 1000 words – more or less!
 - Far more for gigapixel cameras and colour images
 - Less for easily recognizable images
- Thresholds relate to pixel number (and size)
 - Thresholds are scene dependent
 - Thresholds had poor correlations with global image attributes, such as signal-to-noise, contrast, luminance or edge sharpness
 - Images used have small variations in visibility of attributes relative to resolution
- Departures from ‘average’ scenes are due primarily to:
 - visuo-cognitive effects and numerically, in threshold pixel resolution differences

Acknowledgements

Thanks to Fiona Alison for the portrait used in this case study

The author is indebted to the observers who undertook the task in stressful external conditions during periods of lockdown in the UK, caused by the Covid-19 pandemic.

Thanks are due to Professors Sophie Triantaphillidou and Robin Jenkin for their helpful advice and support which is very much appreciated.

Thank You for listening



朋 友

反 苦
(w)

Any Questions or Comments?

中(w)
此: 2

朱明德
2006.6.7.