CS-498 Signals AI

- Themes
  - Much of AI occurs at the signal level
  - Processing data and making inferences rather than logical reasoning
  - Areas such as vision, speech, NLP, robotics
    - methods bleed into other areas (graphics, animation, ...)
  - Linked by the use of statistical tools and ideas from machine learning
  - Much domain knowledge is required to make progress in areas
    - However, they do share tools

Activities

- Mainly lecture
  - but I will require that groups present research papers
- Evaluation
  - 4 projects, which will have a competitive component
    - build a part-of-speech tagger
    - build a face finder
    - build a word spotter
    - build a styleIK (we'll talk about what this means)
  - Final project
    - by choice
  - Participation

Natural Language - Applications

- Machine Translation
  - e.g. South Africa's official languages: English, Afrikaans, the Nguni languages, isiZulu, isiXhosa, isiNdebele, and Siswati, and the Sotho languages, which include Setswana, Sesotho and Sesotho sa Leboa. The remaining two languages are Tshivenda and Xitsonga.
  - български (Bălgarski) - BG - Bulgarian; Čeština - CS - Czech; Dansk - DA - Danish; Deutsch - DE - German; Eesti - ET - Estonian; Elinika - EL - Greek; English - EN; Español - ES - Spanish; Français - FR - French; Gaeilge - GA - Irish; Italiano - IT - Italian; Latviesu valoda - LV - Latvian; Lietuviu kalba - LT - Lithuanian; Magyar - HU - Hungarian; Malti - MT - Maltese; Nederlands - NL - Dutch; Polski - PL - Polish; Português - PT - Portuguese; Română - RO - Romanian; Slovenčina - SK - Slovak; Slovenščina - SL - Slovene; Suomi - FI - Finnish; Svenska - SV - Swedish

More NLP applications

- Question answering
- Information extraction
- Text summarisation
- Information retrieval
- Improved understanding of language, linguistics

Why is NLP hard?

- Meaning is a complex phenomenon
- Sentences are often radically ambiguous

Time flies like an arrow

Fruit flies like a banana

Is this grammatical?

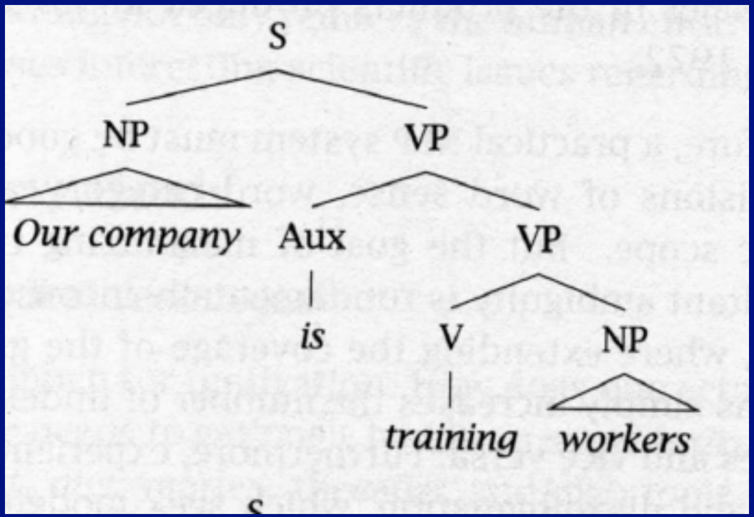
- John I believe Sally said Bill believed Sue saw.
- What did Sally whisper that she had secretly read?
- John wants very much for himself to win.
- Those are the books you should read before it becomes difficult to talk about.
- Those are the books you should read before talking about becomes difficult.
- Who did Jo think said John saw him?
- That a serious discussion could arise here of this topic was quite unexpected.
- The boys read Mary's stories about each other.

Is this grammatical? - Answers

- Y: John I believe Sally said Bill believed Sue saw.
- N: What did Sally whisper that she had secretly read?
- Y: John wants very much for himself to win.
- Y: Those are the books you should read before it becomes difficult to talk about.
- N: Those are the books you should read before talking about becomes difficult.
- Y: Who did Jo think said John saw him?
- Y: That a serious discussion could arise here of this topic was quite unexpected.
- N: The boys read Mary's stories about each other.

Answers due to van Riemsdijk and Williams, 1986, given in Manning and Schutze

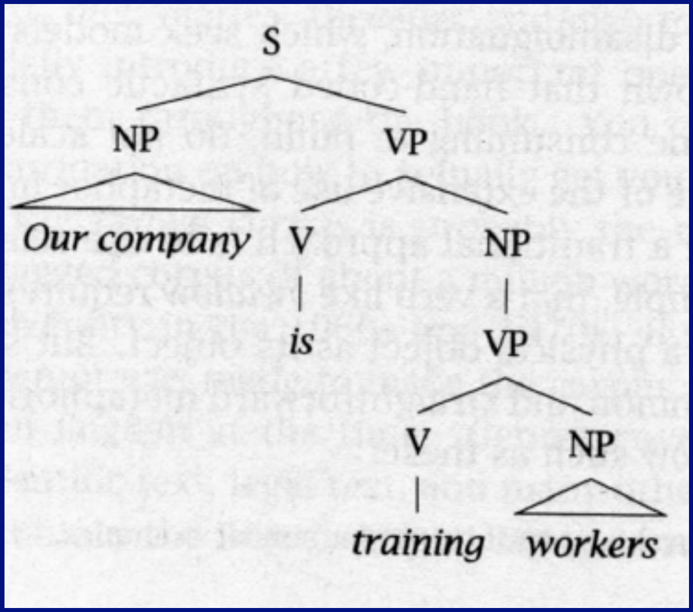
Ambiguity in parsing - 1



- Note:
  - S= sentence
  - NP=noun phrase
  - VP=verb phrase
  - Aux=auxiliary
  - V=verb

Figure from Manning and Schutze

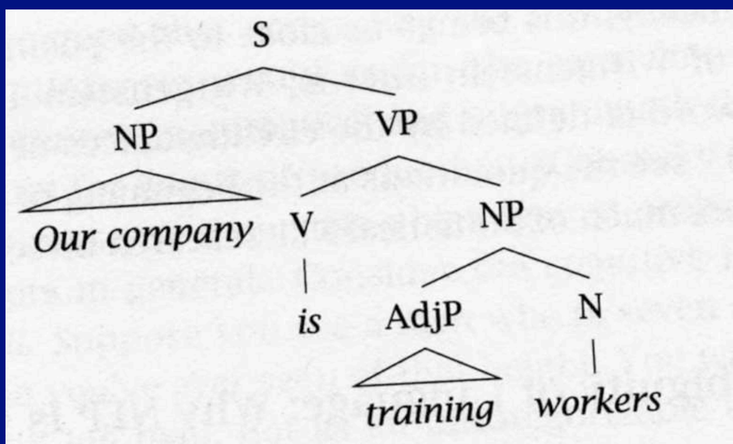
Ambiguity in parsing - 2



- cf
  - Our problem is training workers

Figure from Manning and Schutze

Ambiguity in parsing - 4



- cf
- Those are training wheels

Figure from Manning and Schutze

Ambiguity in parsing -4

- a reasonably sophisticated system gives 455 parses for:
  - List the sales of the products produced in 1973 with the products produced in 1972
- Difficulty:
  - There doesn't seem to be a single grammar that is right
  - choice of grammar is not innocuous:
    - more complex grammars lead to more ambiguous parses
    - less complex grammars can't deal with some (possibly important) special cases

Counts, frequencies and probabilities

- Important phenomena
  - Some things are very frequent
  - Most are very rare
- This is a dominant phenomenon in natural language
  - important in vision, too

What is a word?

- Word token
  - each actual instance of the word
  - There are two “the”s in “the cat sat on the mat”
  - count with multiplicity
  - e.g. 71, 370 in Tom Sawyer
- Word type
  - “the” occurs in “the cat sat on the mat”
  - count without multiplicity
  - e.g. 8, 018 in Tom Sawyer
- Are these two the same word?
  - “ate”, “eat”, “eating”
  - “stock”, “stocking” (perhaps if they’re both verbs, but...)

Word	Freq.	Use
the	3332	determiner (article)
and	2972	conjunction
a	1775	determiner
to	1725	preposition, verbal infinitive marker
of	1440	preposition
was	1161	auxiliary verb
it	1027	(personal/expletive) pronoun
in	906	preposition
that	877	complementizer, demonstrative
he	877	(personal) pronoun
I	783	(personal) pronoun
his	772	(possessive) pronoun
you	686	(personal) pronoun
Tom	679	proper noun
with	642	preposition

Table 1.1 Common words in *Tom Sawyer*.

From Manning and Schutze; recall there are 71, 370 word tokens in Tom Sawyer

Word Frequency	Frequency of Frequency
1	3993
2	1292
3	664
4	410
5	243
6	199
7	172
8	131
9	82
10	91
11-50	540
51-100	99
> 100	102

Table 1.2 Frequency of frequencies of word types in *Tom Sawyer*.

From Manning and Schutze; recall there are 8, 018 word types

Zipf's law

- rank word types by frequency, highest first
- each word type then has:
  - a frequency, f
  - a rank, r
- Zipf's law:
  - $f\,r = \text{constant}$

Word	Freq. (f)	Rank (r)	$f \cdot r$	Word	Freq. (f)	Rank (r)	$f \cdot r$
the	3332	1	3332	turned	51	200	10200
and	2972	2	5944	you'll	30	300	9000
a	1775	3	5235	name	21	400	8400
he	877	10	8770	comes	16	500	8000
but	410	20	8400	group	13	600	7800
be	294	30	8820	lead	11	700	7700
there	222	40	8880	friends	10	800	8000
one	172	50	8600	begin	9	900	8100
about	158	60	9480	family	8	1000	8000
more	138	70	9660	brushed	4	2000	8000
never	124	80	9920	sins	2	3000	6000
Oh	116	90	10440	Could	2	4000	8000
two	104	100	10400	Applausive	1	8000	8000

Table 1.3 Empirical evaluation of Zipf's law on *Tom Sawyer*.

Figure from Manning and Schutze

Zipf's law

- Qualitatively, assuming there are many words
  - few very common words
  - moderate number of medium frequency words
  - very many low frequency words
- Implication:
  - we will spend a lot of effort modelling phenomena we hardly ever observe

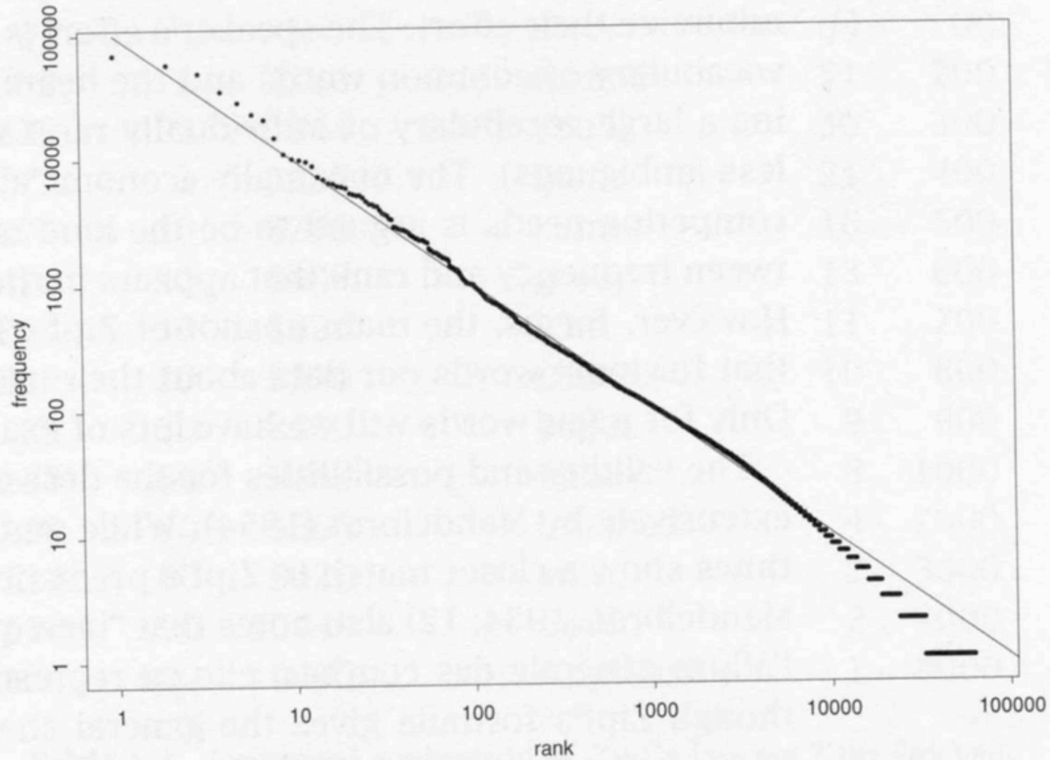


Figure 1.1 Zipf's law. The graph shows rank on the X-axis versus frequency on the Y-axis, using logarithmic scales. The points correspond to the ranks and frequencies of the words in one corpus (the Brown corpus). The line is the relationship between rank and frequency predicted by Zipf for $k = 100,000$, that is $f \times r = 100,000$.

Figure from Manning and Schutze

Collocations: an example

- Collocations are:
  - turn of phrase or accepted usage where whole is perceived to have an existence beyond sum of parts
  - compounds - disk drive
  - phrasal verbs - make up
  - stock phrases - bacon and eggs; steak and kidney; egg and bacon; etc.

Finding possible collocations

- Strategy 1: find pairs of words with high frequency

$C(w^1\ w^2)$	w^1	w^2
80871	of	the
58841	in	the
26430	to	the
21842	on	the
21839	for	the
18568	and	the
16121	that	the
15630	at	the
15494	to	be
13899	in	a
13689	of	a
13361	by	the
13183	with	the
12622	from	the
11428	New	York
10007	he	said
9775	as	a
9231	is	a
8753	has	been
8573	for	a

Table 5.1 Finding Collocations: Raw Frequency. $C(\cdot)$ is the frequency of something in the corpus.

Figure from Manning and Schutze

Finding possible collocations

- Strategy 2: find high frequency pairs of words and then filter them, rejecting any pairs(triples) that do not correspond to part of speech patterns.

Tag Pattern	Example
A N	*linear function*
N N	*regression coefficients*
A A N	*Gaussian random variable*
A N N	*cumulative distribution function*
N A N	*mean squared error*
N N N	*class probability function*
N P N	*degrees of freedom*

Table 5.2 Part of speech tag patterns for collocation filtering. These patterns were used by Justeson and Katz to identify likely collocations among frequently occurring word sequences.

Figure from Manning and Schutze

Finding possible collocations - 3

$C(w^1\ w^2)$	w^1	w^2	Tag Pattern
11487	New	York	A N
7261	United	States	A N
5412	Los	Angeles	N N
3301	last	year	A N
3191	Saudi	Arabia	N N
2699	last	week	A N
2514	vice	president	A N
2378	Persian	Gulf	A N
2161	San	Francisco	N N
2106	President	Bush	N N
2001	Middle	East	A N
1942	Saddam	Hussein	N N
1867	Soviet	Union	A N
1850	White	House	A N
1633	United	Nations	A N
1337	York	City	N N
1328	oil	prices	N N
1210	next	year	A N
1074	chief	executive	A N
1073	real	estate	A N

Table 5.3 Finding Collocations: Justeson and Katz' part-of-speech filter.

Figure from Manning and Schutze

Probability and models

- I assume very basic knowledge of probability and conditional probability.
- Build and investigate procedures to
  - predict words given words
    - e.g. english given french
  - evaluate interpretations of words

Modelling strings of letters

- Alphabet: 27 tokens (each letter, space; no cases)
- Simplest models:
  - M1: tokens are independent, identically distributed, have uniform probability
  - M2: tokens are independent, identically distributed, have different probs.
- Which is better? and why?
- compare P(M1|S) with P(M2|S)
  - using Bayes' rule

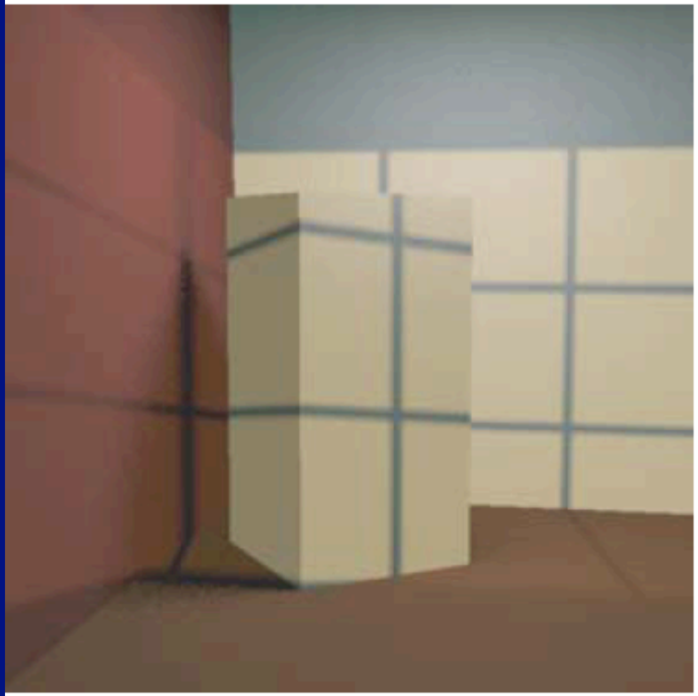
Conditional probability models

- M1 and M2 give quite poor results, M2 much better than M1
- Now consider conditional models
  - we condition a letter on some previous letters
  - 1, 2, ...
  - sometimes known as Markov models
  - these are significantly better in practice
    - we need tools to understand how much better; coming
    - divergence: application of markov models in computer vision

The Importance of Surface Texture

Objects in the real world have rich, detailed surface textures

- to produce believable scenes, we must replicate this detail
- uniformly colored surfaces only get us so far



Generated with Blue Moon Rendering Tools — www.bmrt.org

How Do We Model Intricate Surface Detail?

Approach #1: Explicit geometric representation

- actual polygons that model all the surface variations
- up to some finest level of detail
- may generate a *lot* of polygons

Macroscopic model

Approach #2: Geometry + texture images

- geometry only describes the general shape of the object
- paste an image onto the wall to give the appearance of brick

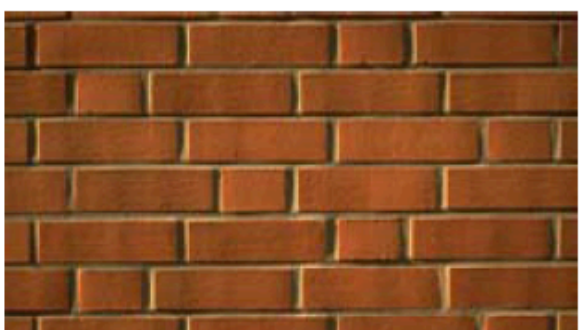
Microscopic model

Often We Use Simple Patterns

Generally useful for skin, bricks, stucco, granite, …

Typically need to repeat texture over the object

- must make sure there are no seams when texture is tiled



Or Given a Model and a Single Texture

Wrap the Texture onto the Model

Sample model from www.cyberware.com

Framework for Texture Mapping

The texture itself is just a 2-D raster image

- acquired from reality, hand-painted, or procedurally generated

Establish a correspondence between surface points & texture

When shading a particular surface point

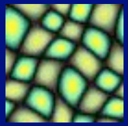
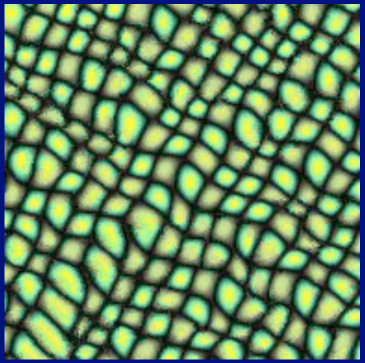
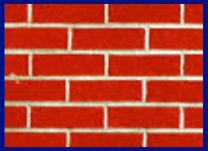
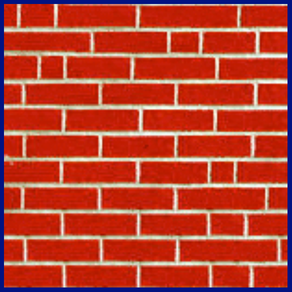
- look up the corresponding pixel in the texture image
- final color of point will be a function of this pixel

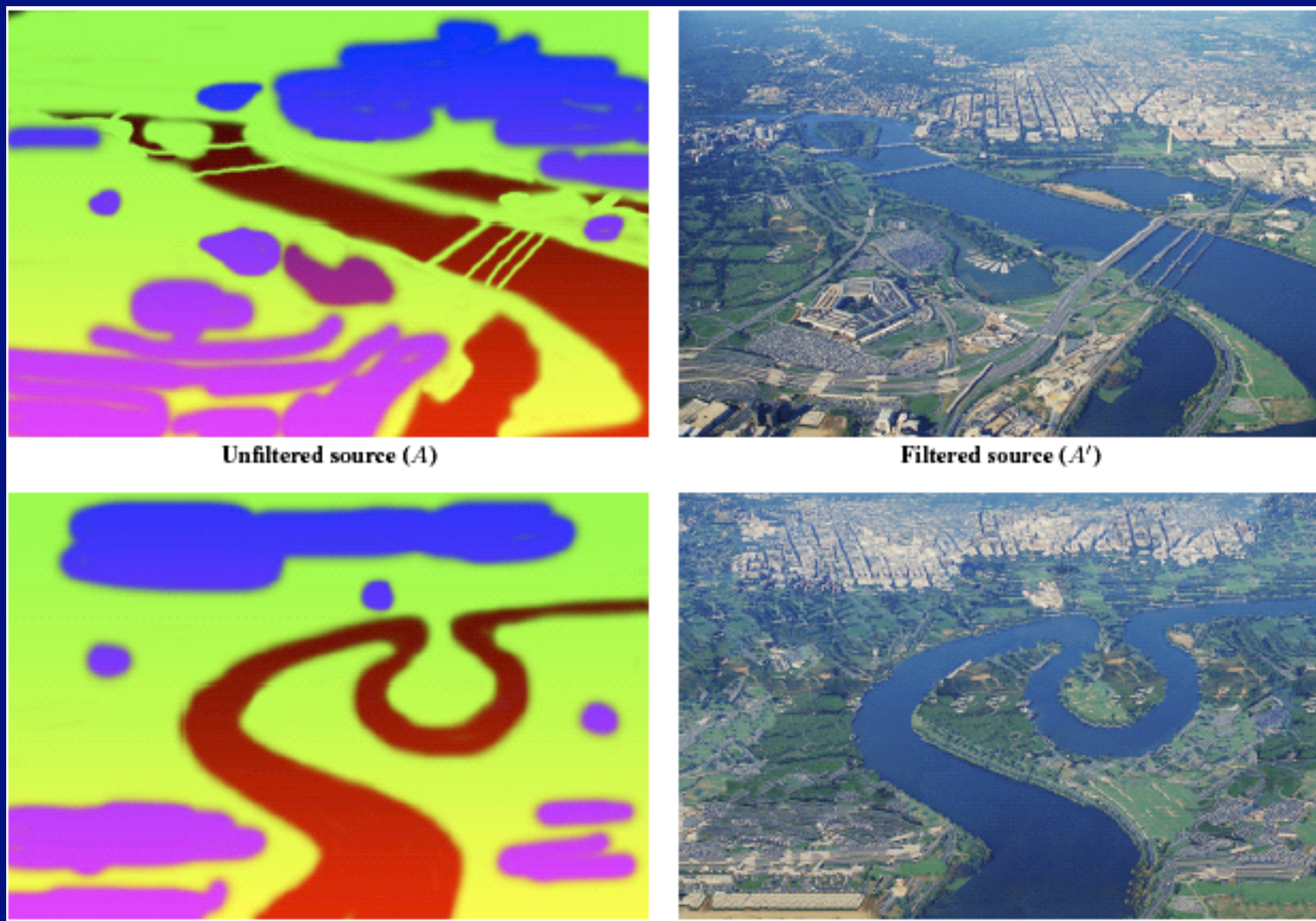
Texture mapping

- Getting enough texture
  - observations
  - buy it
  - tile it
  - synthesize it
- Putting the texture in the right place
  - applying texture to surfaces - artists, parametrization, etc.
  - rendering - ray trace, interpolation

Texture synthesis

- Use image as a source of probability model
- Choose pixel values by matching neighbourhood, then filling in
- Matching process
  - look at pixel differences
  - count only synthesized pixels



From “Image analogies”, Herzmann et al, SIGGRAPH 2001