

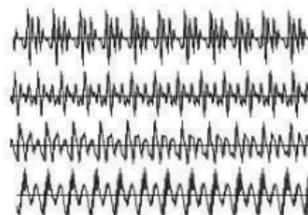
A Text-to-Picture Synthesis System for Augmenting Communication

Xiaojin Zhu Andrew B. Goldberg Mohamed Eldawy
Charles Dyer Bradley Strock

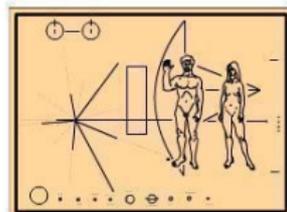
Department of Computer Sciences
University of Wisconsin, Madison, USA

AAAI 2007

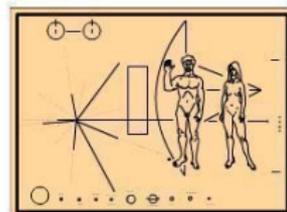
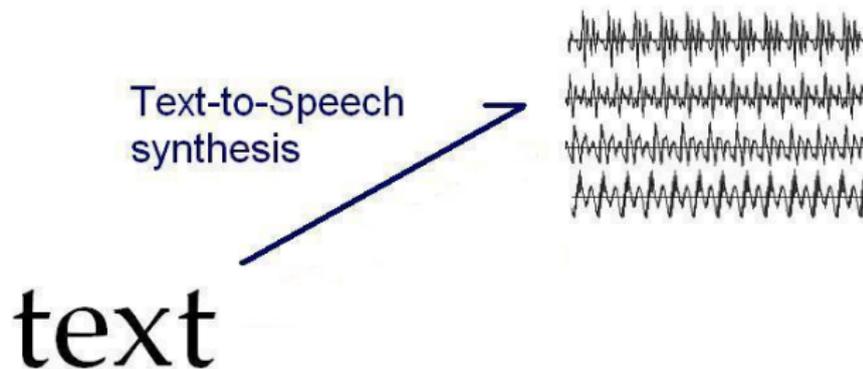
Humans communicate in multiple modalities



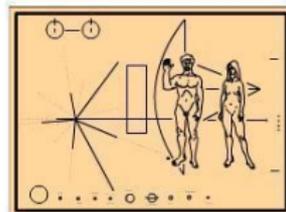
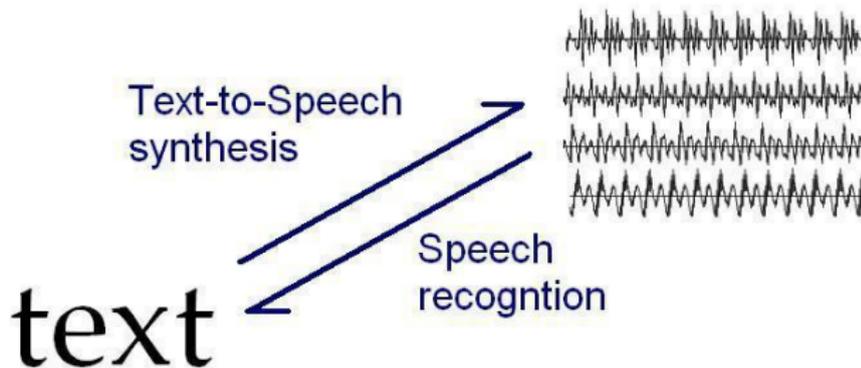
text



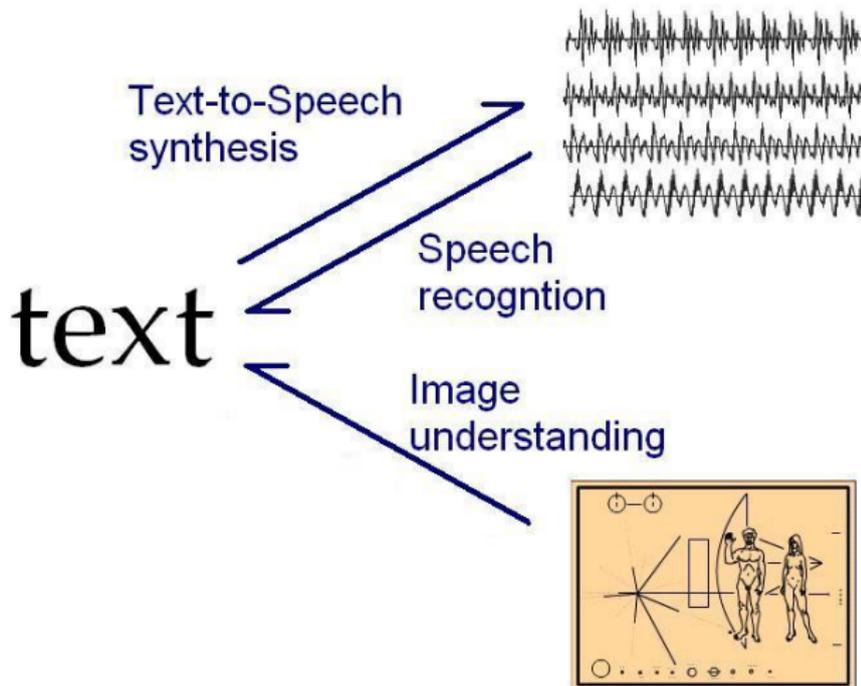
Computer modalities



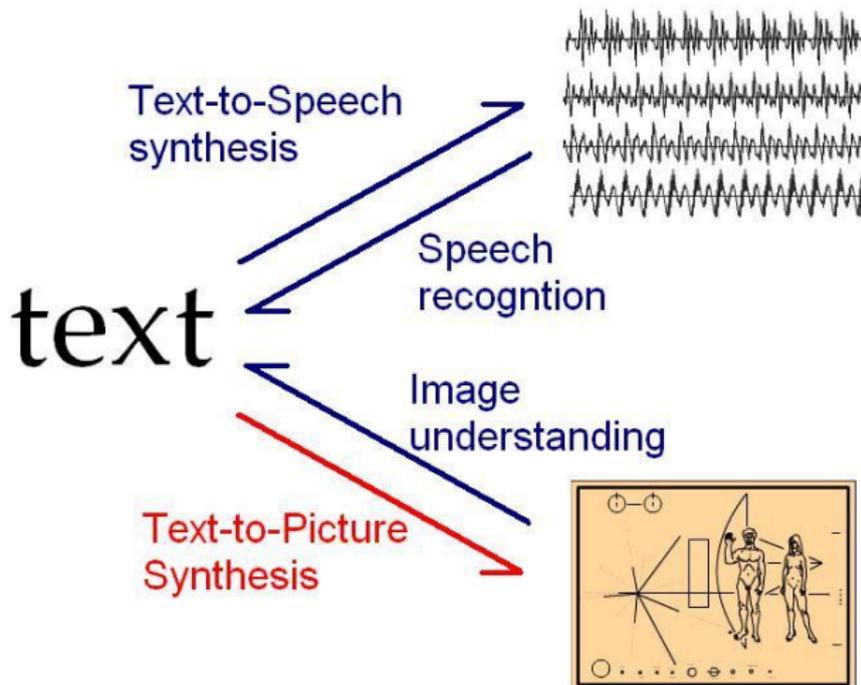
Computer modalities



Computer modalities



Computer modalities



Text-to-Picture (TTP) synthesis

First the farmer gives
hay to the goat. Then
the farmer gets milk from
the cow.

Text-to-Picture (TTP) synthesis

First the farmer gives
hay to the goat. Then
the farmer gets milk from
the cow.



Text-to-Picture (TTP) synthesis

First the farmer gives hay to the goat. Then the farmer gets milk from the cow.



Goal

To convey the gist of general, unrestricted text.

Outline

- ① Discussing TTP as a research field.
- ② Our first attempt at a general purpose TTP system.

Applications of Text-to-Picture

- Literacy development: young children, 2nd language speakers

Applications of Text-to-Picture

- Literacy development: young children, 2nd language speakers
- Assistive devices: learning disability, traumatic brain injury

Applications of Text-to-Picture

- Literacy development: young children, 2nd language speakers
- Assistive devices: learning disability, traumatic brain injury
- Universal language

Applications of Text-to-Picture

- Literacy development: young children, 2nd language speakers
- Assistive devices: learning disability, traumatic brain injury
- Universal language
- Document summarization

Applications of Text-to-Picture

- Literacy development: young children, 2nd language speakers
- Assistive devices: learning disability, traumatic brain injury
- Universal language
- Document summarization
- Image authoring tool

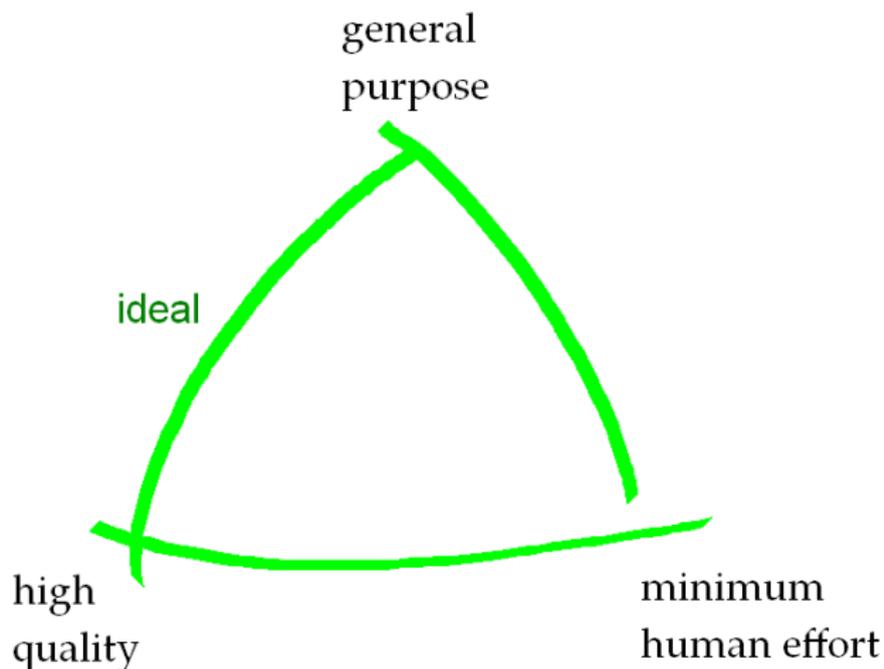
Three qualities of a Text-to-Picture system

general
purpose

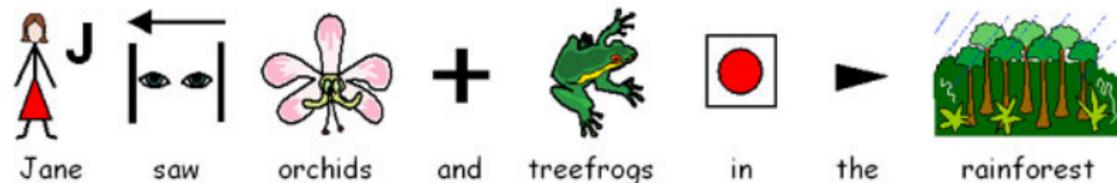
high
quality

minimum
human effort

Three qualities of a Text-to-Picture system



Prior work 1: "Writing with Symbols"



- Rebus symbols (www.widgit.com)
- Writing with Symbols (www.mayer-johnson.com)

Prior work 1: “Writing with Symbols”

general
purpose

Writing-with-Symbols

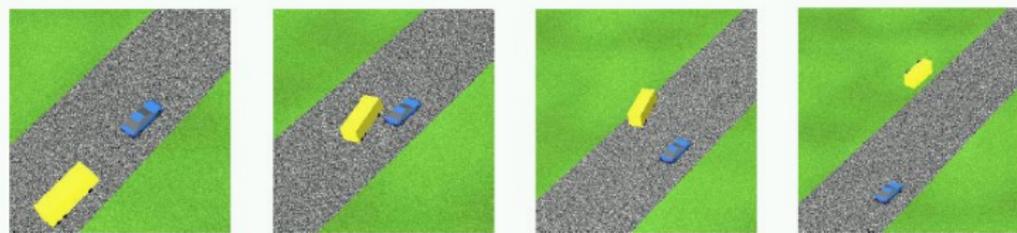


high
quality

minimum
human effort

Prior work 2: CarSim

A bus accident in southern Afghanistan last Thursday claimed 20 victims. Additionally, 39 people were injured in the accident, which occurred early Thursday morning twenty kilometers north of the city Kandahar. The bus was on its way from Kandahar towards the capital Kabul when it left the road while overtaking and overturned, said general Salim Khan, assistant head of police in Kandahar. The state of some of the injured was said to be critical.



[Johansson, Berglund, Danielsson and Nugues. IJCAI 2005]

Prior work 2: CarSim

general
purpose

high
quality

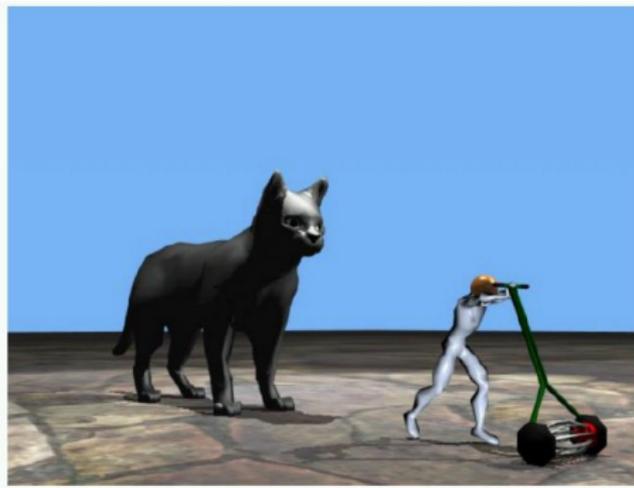


Writing-with-Symbols

minimum
human effort

Prior work 3: WordsEye

The lawn mower is 5 feet tall. John pushes the lawn mower.
The cat is 5 feet behind John. The cat is 10 feet tall.



[Coyné and Sproat. SIGGRAPH 2001] (www.wordseye.com)

Prior work 3: WordsEye

general
purpose



Goal of this work

To convey the gist of general, unrestricted text.



Approaches to Text-to-Picture

- 1 “Canned” pictures
- 2 Model-based
- 3 Concatenative

Components of our TTP system

- ① Keyphrase selection
- ② Image unit selection
- ③ Image layout
- ④ Evaluation

Step 1: keyphrase selection

Problem: decide which words to draw.

Step 1: keyphrase selection

Problem: decide which words to draw.

- TextRank keyword summarization [Mihalcea and Tarau 2004].

Step 1: keyphrase selection

Problem: decide which words to draw.

- TextRank keyword summarization [Mihalcea and Tarau 2004].
- Graph on nouns, proper nouns, adjectives.

Step 1: keyphrase selection

Problem: decide which words to draw.

- TextRank keyword summarization [Mihalcea and Tarau 2004].
- Graph on nouns, proper nouns, adjectives.
- Edges for word co-occurrence.

Step 1: keyphrase selection

Problem: decide which words to draw.

- TextRank keyword summarization [Mihalcea and Tarau 2004].
- Graph on nouns, proper nouns, adjectives.
- Edges for word co-occurrence.
- Random walk on the graph.

Step 1: keyphrase selection

Problem: decide which words to draw.

- TextRank keyword summarization [Mihalcea and Tarau 2004].
- Graph on nouns, proper nouns, adjectives.
- Edges for word co-occurrence.
- Random walk on the graph.
- Stationary distribution (PageRank) as word importance.

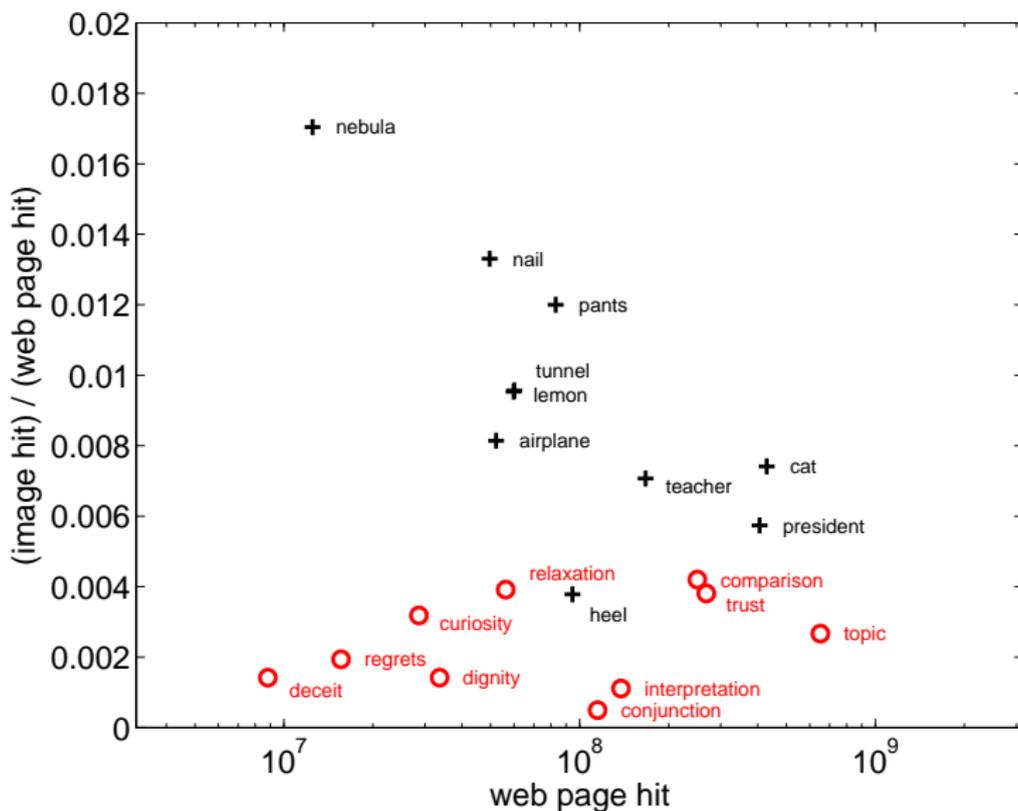
Step 1: keyphrase selection

Problem: decide which words to draw.

- TextRank keyword summarization [Mihalcea and Tarau 2004].
- Graph on nouns, proper nouns, adjectives.
- Edges for word co-occurrence.
- Random walk on the graph.
- Stationary distribution (PageRank) as word importance.
- Important difference: word **picturability** used for teleporting probability.

Word picturability

We want to select picturable words.



Word picturability model

- 253 candidate features from Google, Yahoo!, Flickr search
- Best feature: $x = \log(\text{Google image hits}/\text{Google Web page hits})$
- Logistic regression
- $\Pr(\text{picturable}|x) = 1 / (1 + \exp(-2.78x - 15.4))$

Step 2: Image unit selection

Problem: find the best image for a word.

Step 2: Image unit selection

Problem: find the best image for a word.

- Collect top image search results



Step 2: Image unit selection

Problem: find the best image for a word.

- Collect top image search results



- Segmentation

Step 2: Image unit selection

Problem: find the best image for a word.

- Collect top image search results



- Segmentation
- Cluster image segments by color

Step 2: Image unit selection

Problem: find the best image for a word.

- Collect top image search results



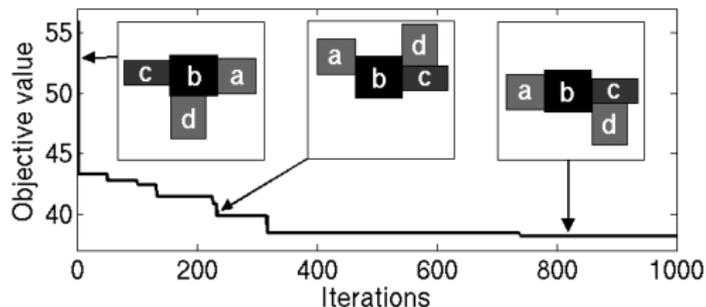
- Segmentation
- Cluster image segments by color
- Select the image containing the segment at the center of the largest cluster.

Step 3: Image layout

Problem: put the images together.

- minimum overlap
- important words at center
- close in text, close in picture

Optimization.



Evaluation

The large chocolate-colored horse trotted in the pasture.

Evaluation

The large chocolate-colored horse trotted in the pasture.



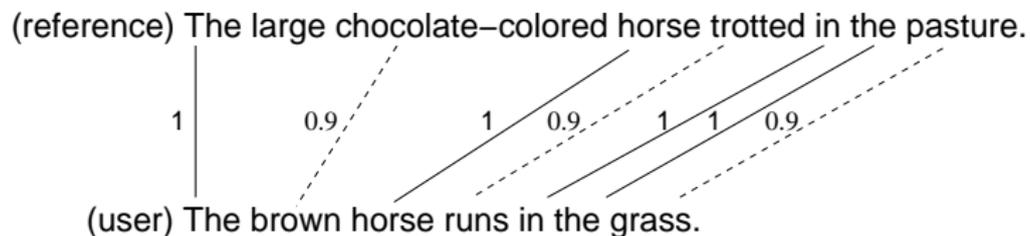
Evaluation

The large chocolate-colored horse trotted in the pasture.



The brown horse runs in the grass.

Evaluation

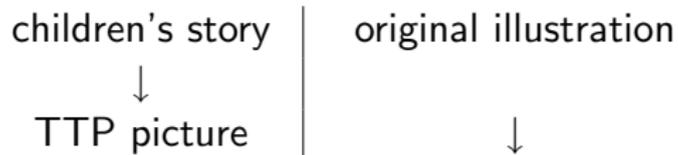


- synonyms
- greedy word alignment
- F-measure from precision and recall

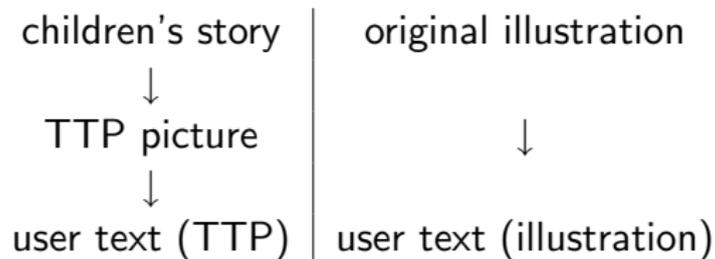
Experiment on children's books

children's story | original illustration

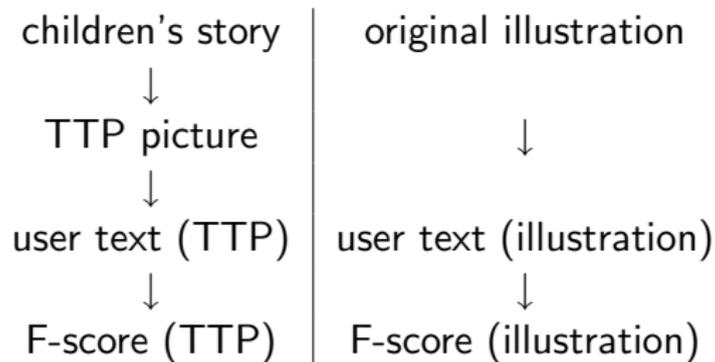
Experiment on children's books



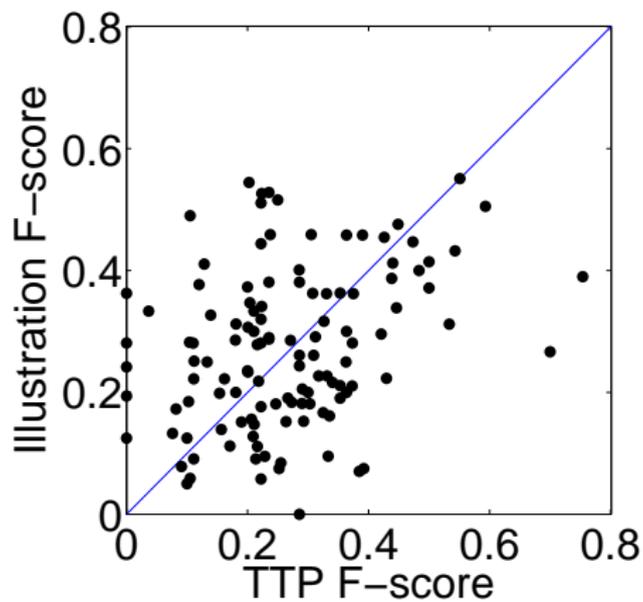
Experiment on children's books



Experiment on children's books

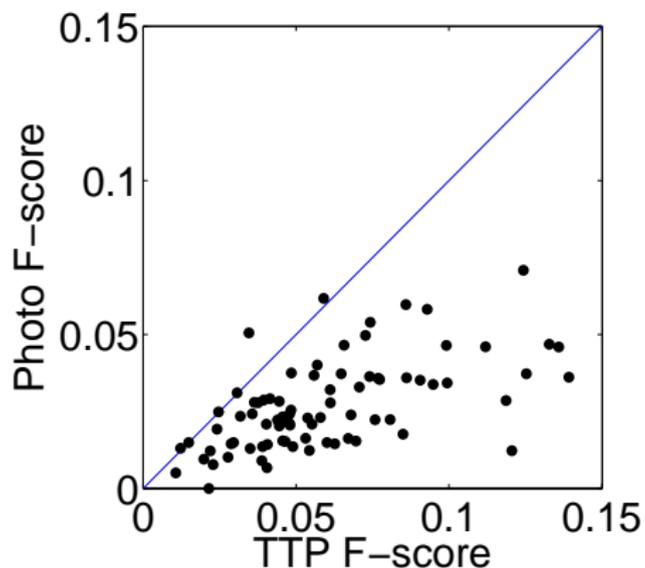


Experiment on children's books



Experiment on news articles

Original news photo vs. TTP generated picture + news photo



Our system is far from perfect

The girl loved the dog. The girl loved the dog's soft eyes and warm nose and big paws. The girl wished she had a dog.



“A girl’s pet puts its paw on her nose.”

“The dog walked up to the girl and sniffed her.”

“The dog bit the girl in her nose and ran away.”

“The girl’s nose smelled the dog and monkey as they walked away.”

“The girl walked her dog and saw a hairy man with a big nose.”

“The girl monkey nose smells dog paw prints.”

Conclusions

- 1 Text-to-Picture is a research area which deserves more attention.
- 2 We have a preliminary system. Much still to be done.

Conclusions

- 1 Text-to-Picture is a research area which deserves more attention.
- 2 We have a preliminary system. Much still to be done.

Thank you.