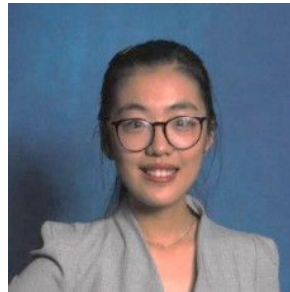


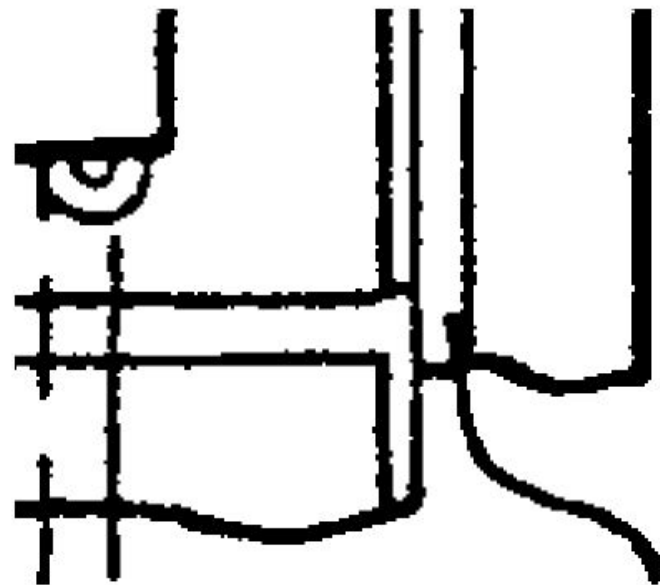
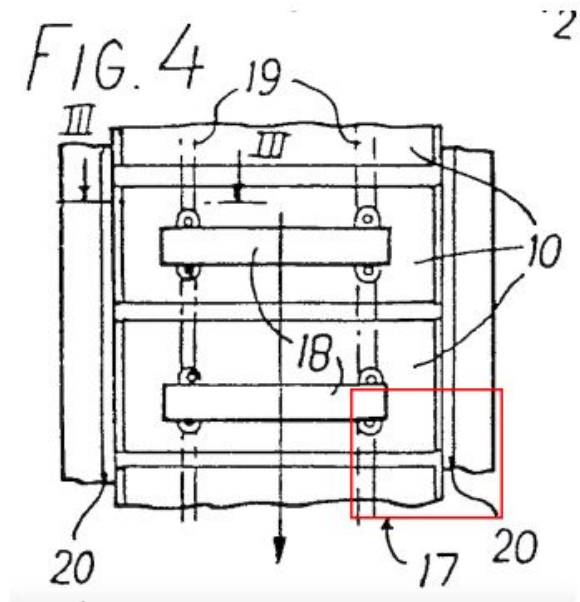
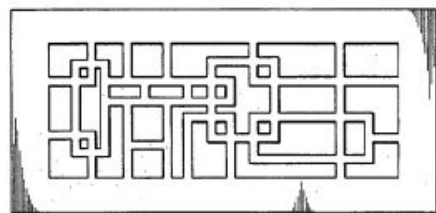
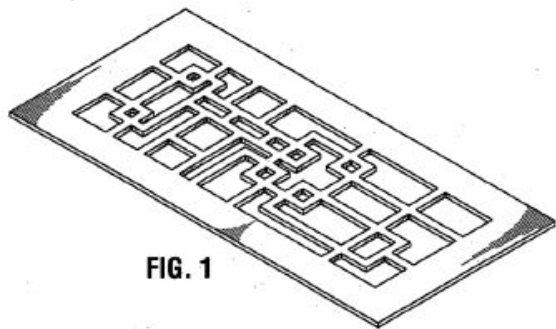
Diagram Image Retrieval and Analysis: *Challenges and Opportunities*

[Liping Yang](#), [Ming Gong](#), [Vijayan K. Asari](#)



Poster presented at CVPR 2020 [DIRA workshop](#)

Introduction and motivation



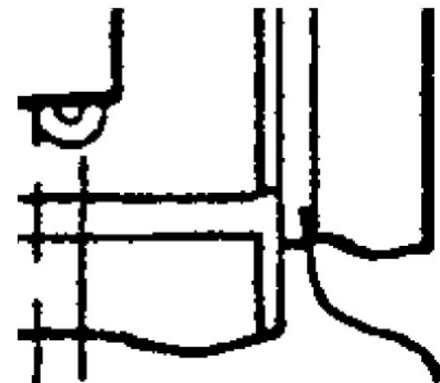
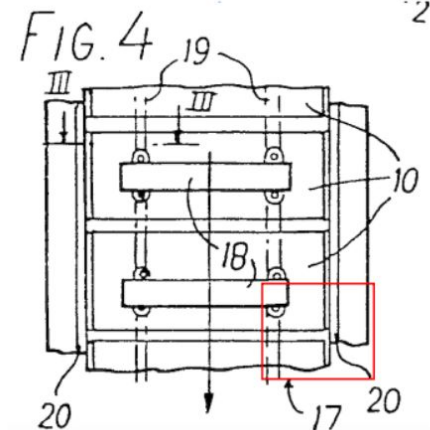
(a) An industrial designs diagram image

(b) A patent diagram image

(c) A zoomed-in detail from (b)

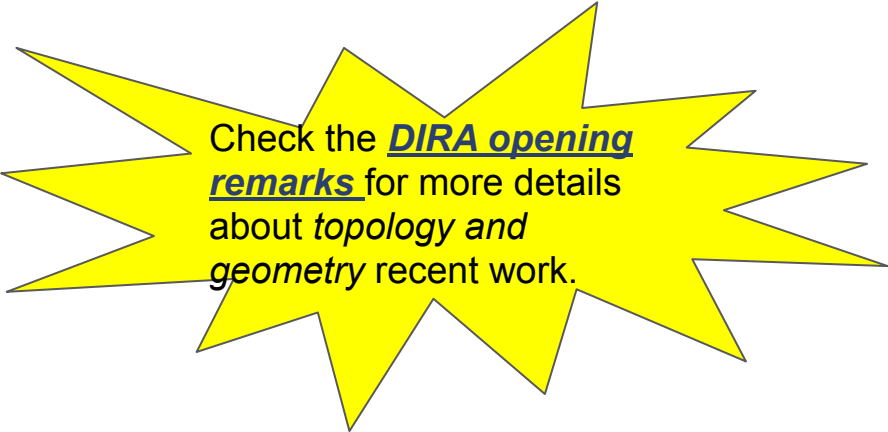
Introduction and motivation

- DIRA Research is much less well developed
 - Dominant features in scientific diagram images are (complex) **shape and topology**,
 - *NO color and intensity* features and
 - *little texture* information
 - Zigzag noises are a bottleneck for DIRA using existing methods
 - Diagram images often contain text annotations
- Draw attention of CV researchers and practitioners



Methods reviewed

- CBIR-based methods
- Texture-based methods
- Shape-based methods
- **Topology and geometry-based approaches**

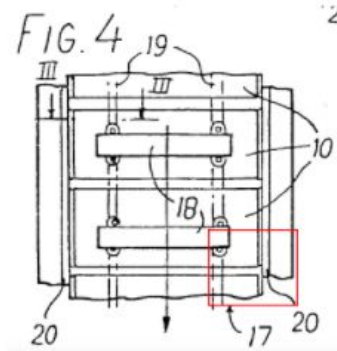


Check the [DIRA opening remarks](#) for more details about *topology and geometry* recent work.

Challenge illustration: *CBIR-based*

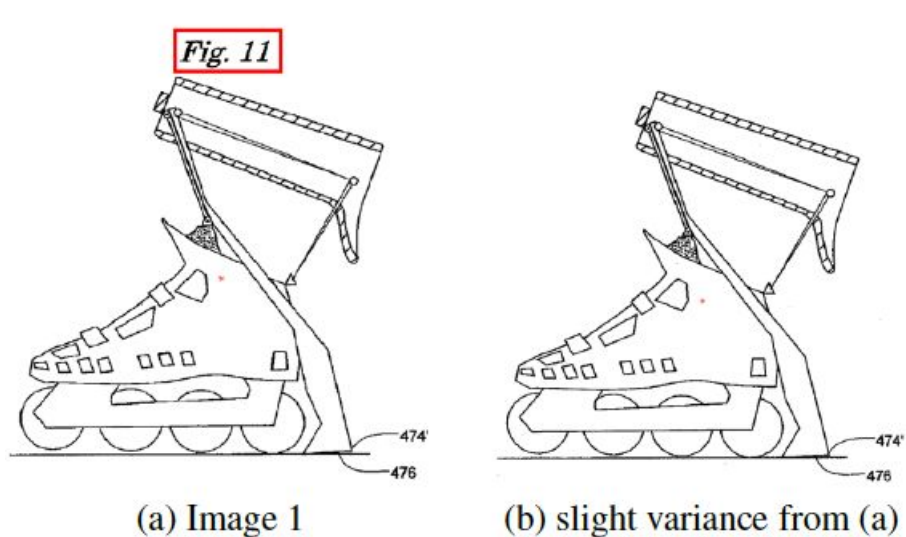
- Adaptive hierarchical density histogram (AHDH)
- AHDH downsides:
 - not rotation invariant,
 - lack of sub-image matching capability,
 - very sensitive to text annotations (very common in diagram images)
 - **not robust**

See our paper **appendix** for tech details about how *AHDH* works (with illustrative examples)

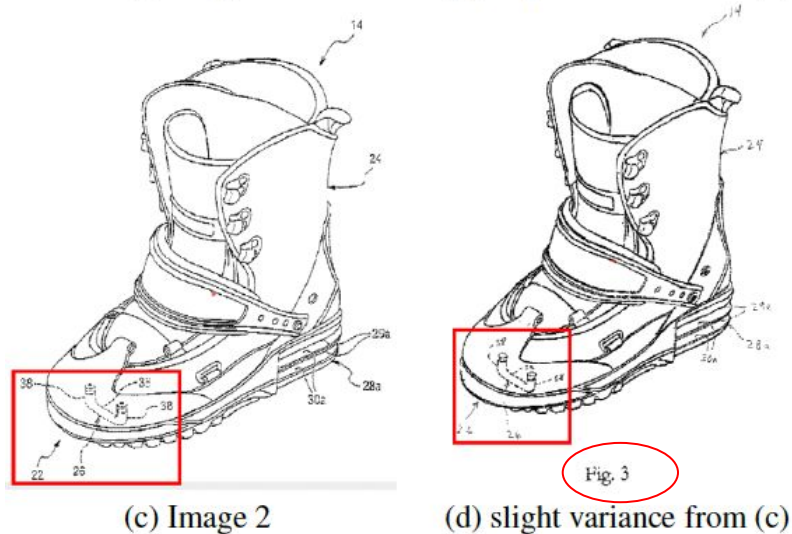


Challenge illustration: *CBIR*-based

Not a difficult task for humans to tell the two pairs of images are very visually similar but not so for machines.



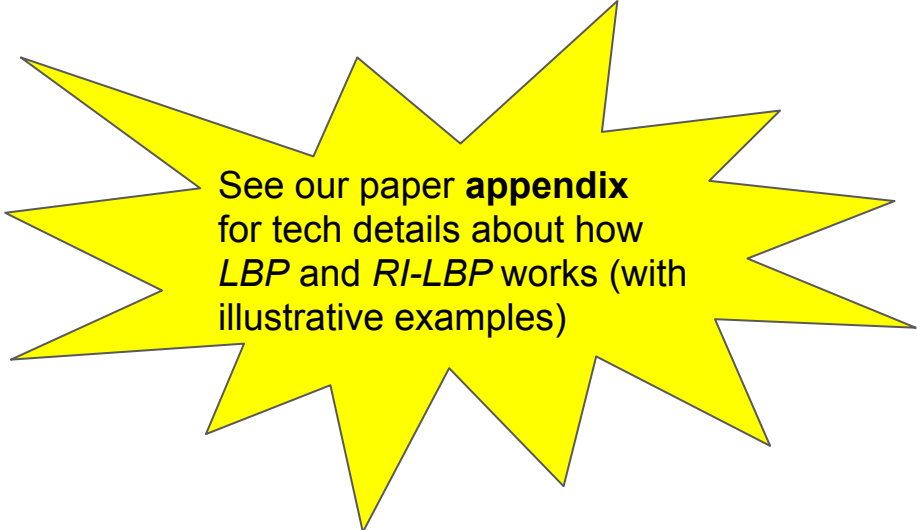
$Edist [(a), (b)]$ is 4.2243



$Edist [(c), (d)]$ is 12.0803

Challenge illustration: *Texture-based*

- Local binary patterns (LBP) -- *simple and powerful but not rotation invariant*
- **Rotation invariant LBP (RI-LBP)**
- **Do not use texture-based features alone for DIRA**, combine it with other features, such as shape descriptors



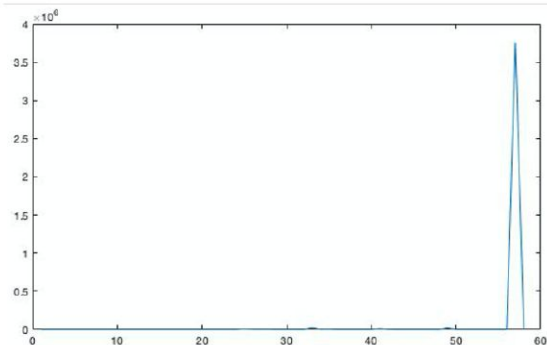
See our paper **appendix** for tech details about how *LBP* and *RI-LBP* works (with illustrative examples)

Challenge illustration: *Texture-based*

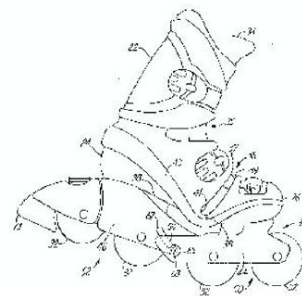
RI-LBP indeed rotation invariant, including for diagram images.



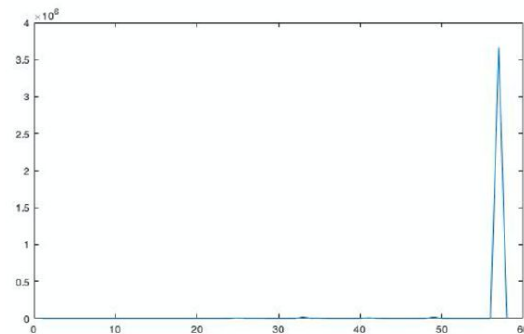
(a) Original Image



(b) RI-LBP histogram of (a)



(c) Rotated (a)

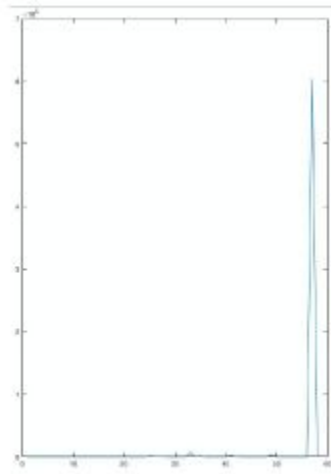
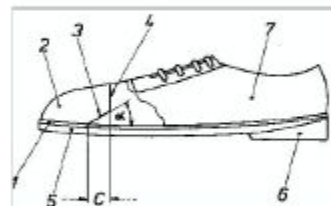
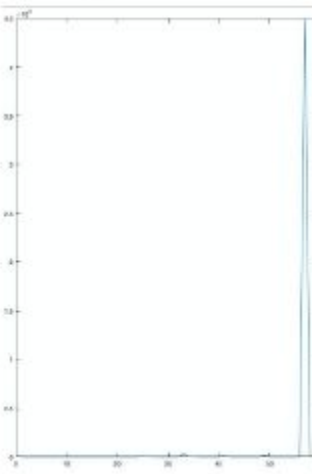
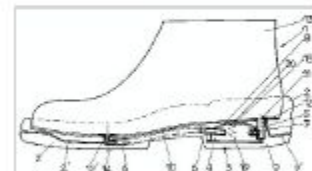
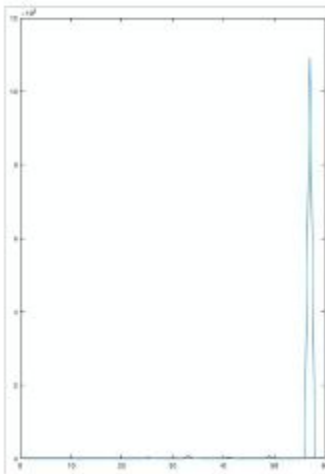
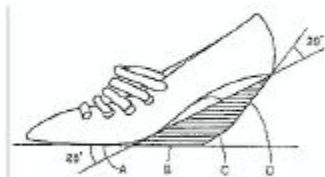
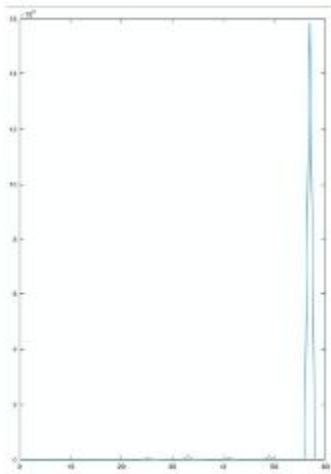
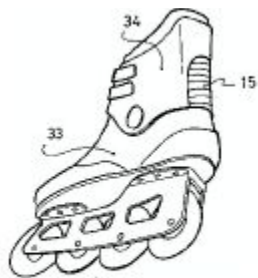


(d) RI-LBP histogram of (c)

Challenge illustration: *Texture-based*

RI-LBP cannot tell visual similarity between diagram images.

The peak occurs at the same position with different amplitude.



Challenge illustration: *Shape-based*

- Shape context (SC)
- Region-based shape descriptor (RBSD)
- **Neither SC nor RBSD can tell and rank the visual similarity of diagram images well.** *The label and annotation text really confuse shape-based descriptors.*

Challenge illustration: SC-based

Image 1	Image 2	Shape extracted from image 1	Shape extracted from image 2	Matching	Std diff	Cos diff
					20.01	0.40
					36.93	0.77
					22.04	0.48

Std diff:
Pearson's
chi-squared test
(used to count cost
matrix)

Cos diff:
cosine distance

Apparently
NOT
visually
similar

Challenge illustration: *RBSD*-based

Image 1	Image 2	E_A	E_V	E_H	E_Z	E_{RBSD}
		0.27	1.26	1.26	0.29	3.08
		0.07	0.65	0.65	0.18	1.55
		0.30	0.93	0.93	0.19	2.35

E_A : Edist of **angular** radial histogram

E_V : Edist of **vertical** histogram

E_H : Edist of **horizontal** histogram

E_Z : Edist of **Zernike** moment

$$E_{RBSD} = E_A + E_V + E_H + E_Z$$

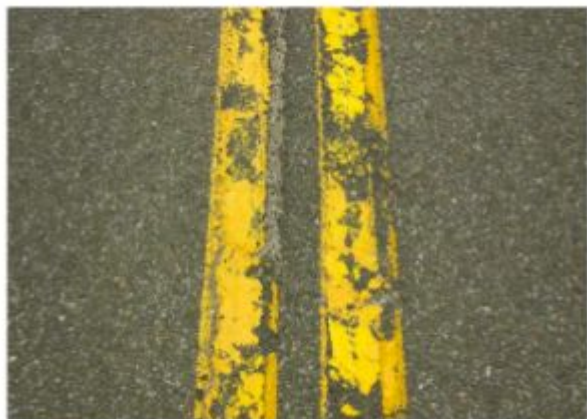
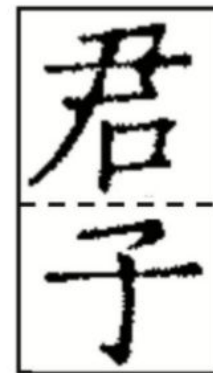
E_{RBSD} values cannot effectively and accurately tell the visual similarity among the three pairs of patent diagram images.

DIRA Tech Opportunities:

- General directions
 - *new image representations, methods and algorithms, and similarity metrics,*
 - *psychologically inspired conceptual schemes and theories*
- Large representative benchmark data sets
- Topology related opportunities (*Check [DIRA opening remarks](#)*)
- Specific opportunities
 - *RI-AHDH*
 - *Integration of RI-LBP*
 - *Extension of SC*
 - *Extension of RBSD*

DIRA Application Opportunities

- Patent image retrieval
- OCR and text recognition
- Road lane line detection for autonomous driving



Check my ***DIRA opening remarks***

(talk video & slides)

for our recent novel DIRA work

cvpr-dira.lipingyang.org/schedule





Check our DIRA paper for more details at
<http://cvpr-dira.lipingyang.org/accepted-papers>

Please email lipingyang@unm.edu if any questions

