POWER PAPERS

Some Practical Pointers (Part 2)

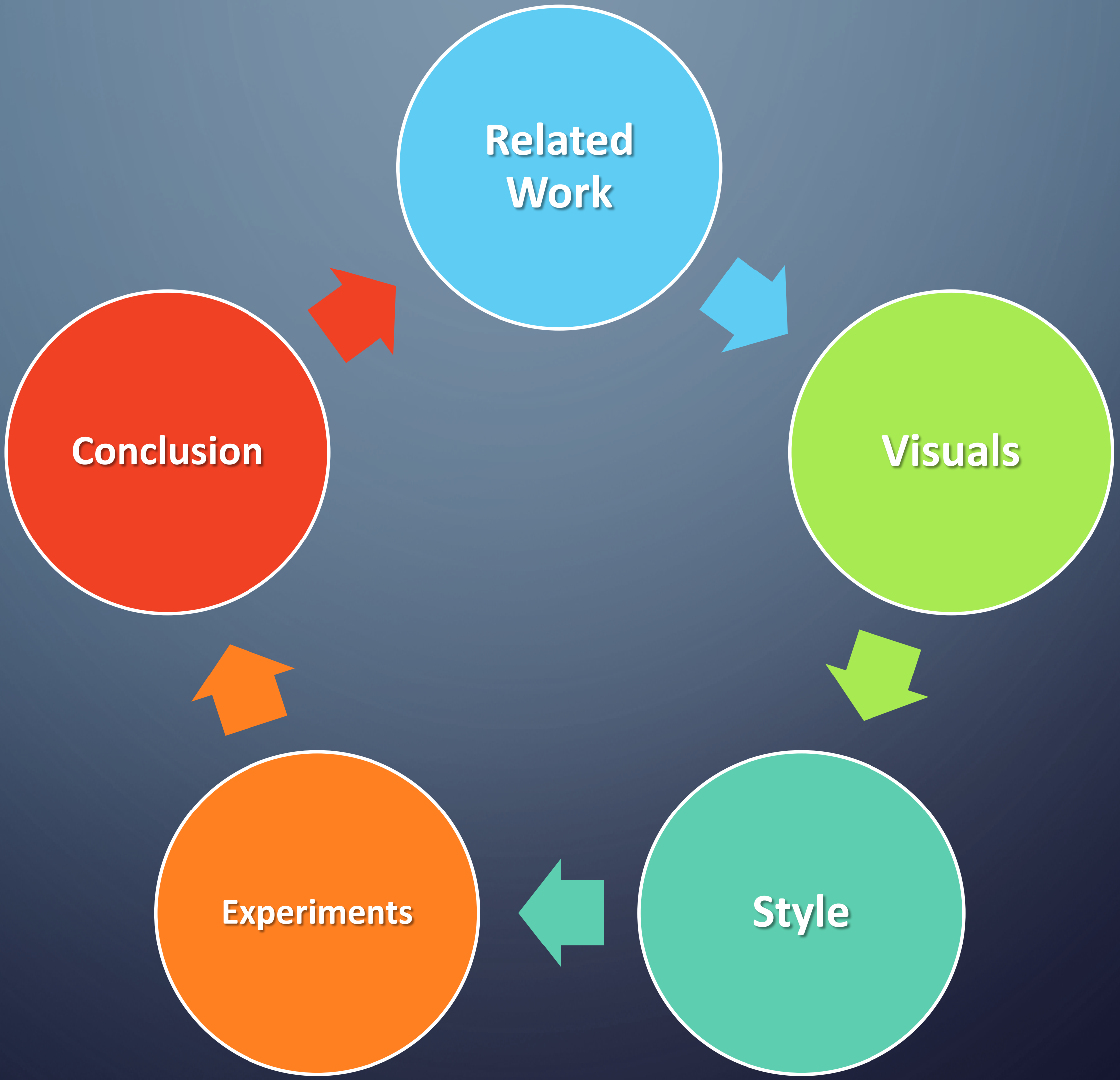
Dr. Terence Sim
School of Computing
National University of Singapore

4 Aug 2021, 12 Feb 2020, 1 Mar 2019, 17 Mar 2018, 18 Apr 2017

PREVIOUSLY …

We talked about the Introduction and Title.

Now,



ASSUMPTIONS

You agree that writing well is important.

You (can) write grammatically correct sentences.

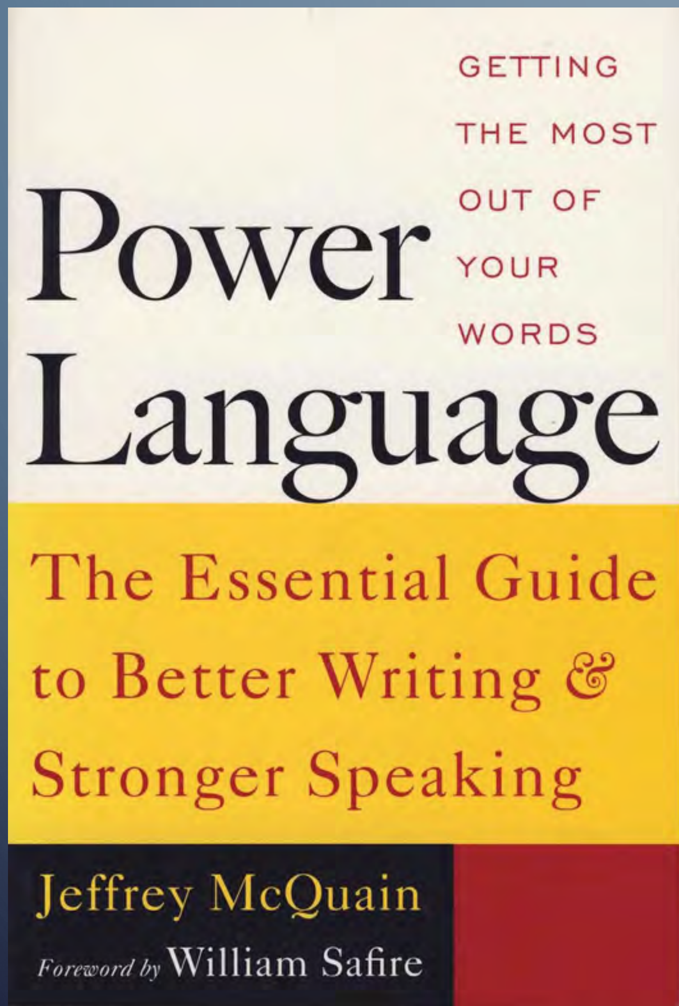
- Occasional mistakes are ok.

You have good research to write about.

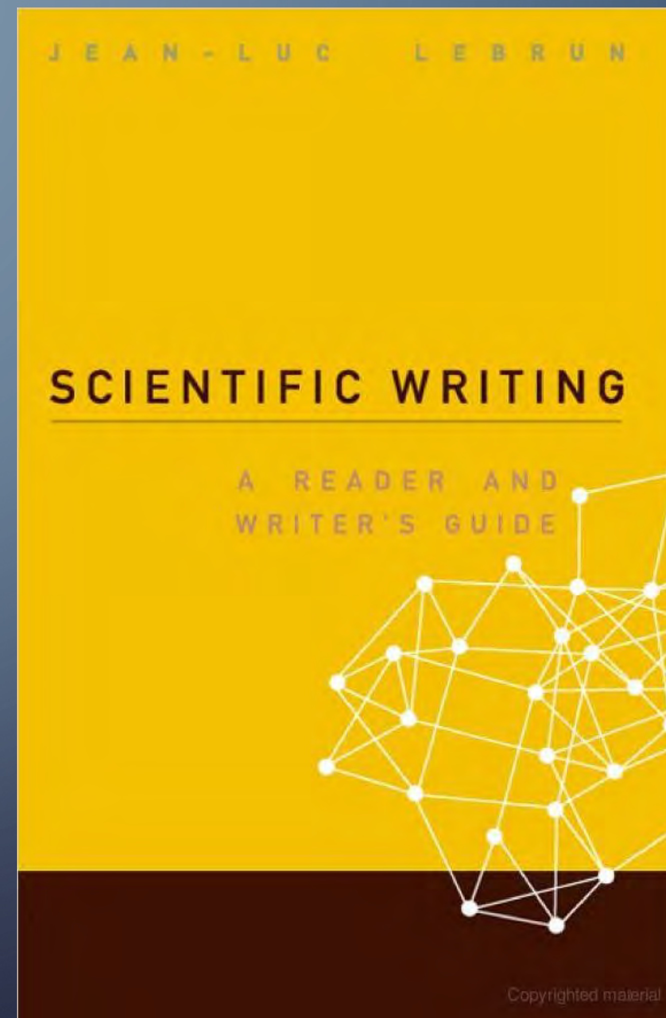
BOOKS

www.scientific-writing.com

[1]



[2]



Set & Manage your Reader's

Anticipate what your reader is thinking, and guide it!

EXAMPLE

"Our data reveal that, contrary to Tom Smith's assumption (4), the pinhole corrosion byproducts do migrate to form part of the top layer material."

Example taken from Lebrun [2]

WHAT ABOUT THIS?

"Our data reveal that the pinhole corrosion byproducts migrate to become part of the top layer material. These findings contradict Tom Smith's assumption (4)."

Example taken from Lebrun [2]

Your words create expectations in the reader’s mind

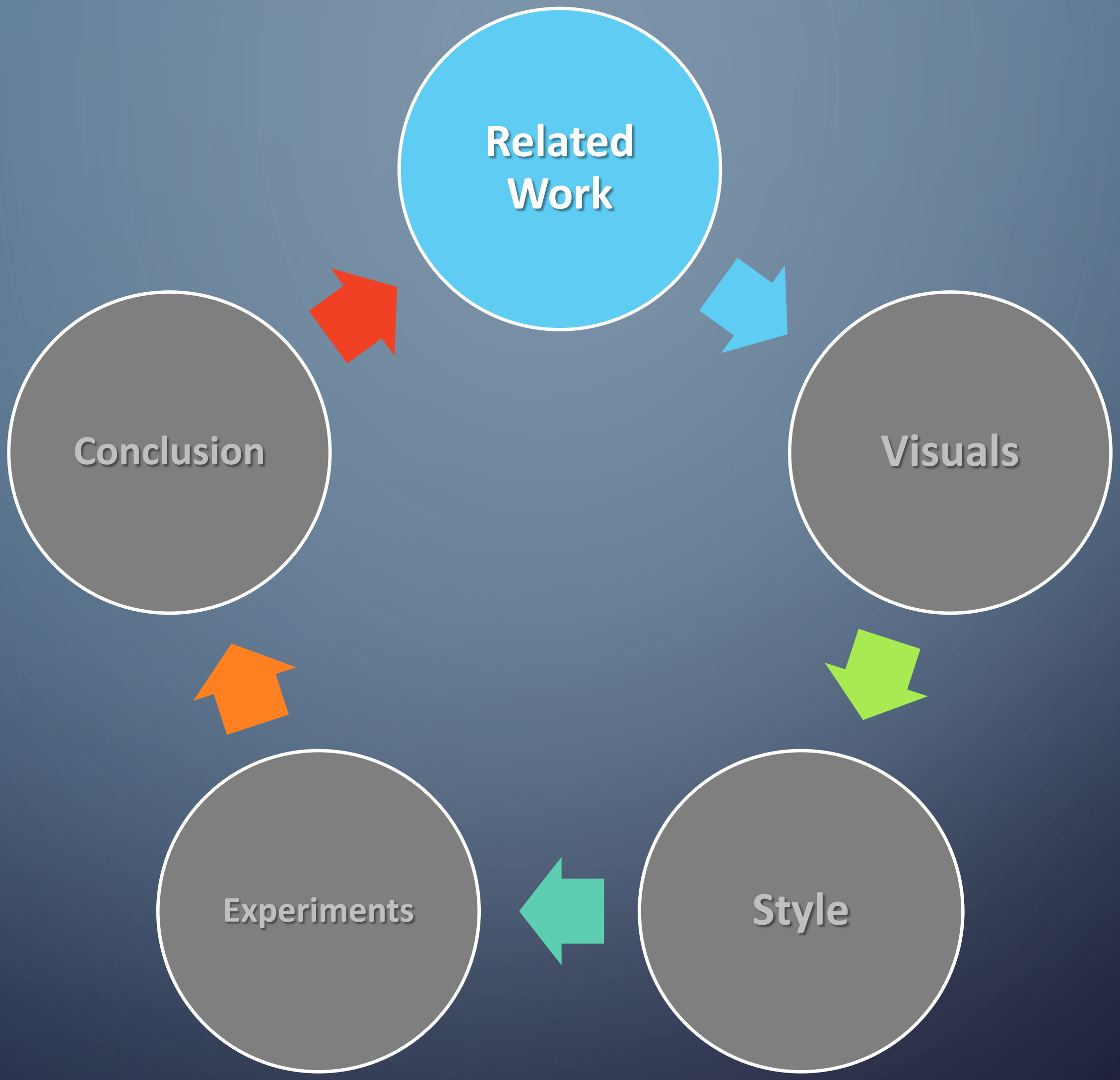
Choose

Fulfill expectations

- Happy reader

Dash expectations

- Frustrated reader


Related Work
Visuals
Style
Experiments
Conclusion

ANNOTATED GOOGLE LISTING

S.J. Shepherd was the first to investigate on Continuous Keystroke Authentication [1] using mean and the standard deviation of Held Times and Interkey Times.   Villani et al., conducted studies on Keystroke Biometric in Long-Text input under Application-Oriented conditions [7]. Keystroke Analysis of Different Languages was conducted by Gunetti et al., [8] which emphasis that Keystrokes can be used as a Biometric in a Language independent setting.

Example taken from Janakiraman and Sim [3]

COMPARE AND CONTRAST

In our literature search, we note that S.J. Shepherd [1] was perhaps the first to explore using Keystroke Dynamics for continuous authentication, using the rate of typing. The system authenticated the user based only on the mean and standard deviation of the Held Times and the Interkey Times, irrespective of the key being pressed. Although it worked for a user population of four, the accuracy of the system is likely to decrease as the number of users increase. There is no guarantee that these features are sufficiently discriminative. Indeed, our experiments conducted with a larger pool of 22 users confirm this.

Example taken from Janakiraman and Sim [3]

USE A TABLE

Table 2.2: A comparison of methods: flash & no-flash, multiple flash, and our selective re-flashing approach.

	Flash and No-flash	Multiple Flash	Our Method
Input	A flash image and an ambient image	Multiple flash images with known flash intensity	Two flash images, only the ratio of two flash intensity is required
Static scene	Strongly dependent	Strongly dependent	Weakly dependent
Method	• *Joint bilateral filter*: transferring texture or color from flash to ambient • *Gradient projection*: remove reflection or hot spots	• *Linear model*: recovering and re-rendering the ambient image	• *Gradient decomposition*: recovering the ambient and flash-only image, selective re-flashing
Trade-off	Parameter setting	Calibration	Calibration
Artifacts handling	• Shadows, specularities, and reflections are detected and removed using different ad hoc methods.	• Shadows and interreflection can be well separated. • Specularities are removed using two flashes.	• Shadows and inter-refelction are naturally separated and selectively suppressed. • Specularities can be effectively detected and removed based on visual cues from two flash images.
Visual quality	• Enhancing the image quality by fusing ambient and flash images, or removing flash artifacts. • The final result is largely dependent on the visual quality of captured ambient image.	• Re-rendering various effects. • But recovering is sensitive to noise.	• Recovering the ambient and flash-only images with high visual quality. • Allowing user to selectively re-flash or keep the ambiance of desired regions.

Example taken from Xiaopeng Zhang [4]

CATEGORIZE BY PROBLEM TYPE, NOT BY METHODS

Blur Type \ Input		Single Image	Multiple Image	Image Sequence	Other Modalities
Motion blurring	2D Motion	[24, 68, 10, 13, 62, 43, 41, 42, 61, 8, 3, 17, 49, 50, 63, 51, 47, 37, 38, 67, 20, 14, 54, 19]	[60, 32]	[7, 5]	[11, 12, 57, 48, 35, 16, 36, 29, 69, 34]
	3D Motion				[11, 12, 57, 48, 35, 16, 36, 29, 69, 34]
Incorrect Focus		[24, 68, 44, 10, 13, 62, 43, 41, 42, 61, 8, 3, 17, 49, 50, 63, 51, 47, 37, 38, 67, 20, 14, 54]			
Large aperture					
Optical imperfection					
Weather			[52]	[22]	
Image sensor	Blooming				
	Low resolution	[2, 4]	[4]	[5, 6, 58, 15, 4]	
Image processing	Bayer Pattern				[27, 46]
	Dynamic Range		[18, 53]		[53]
	Image Compression	*Not covered in this paper*			

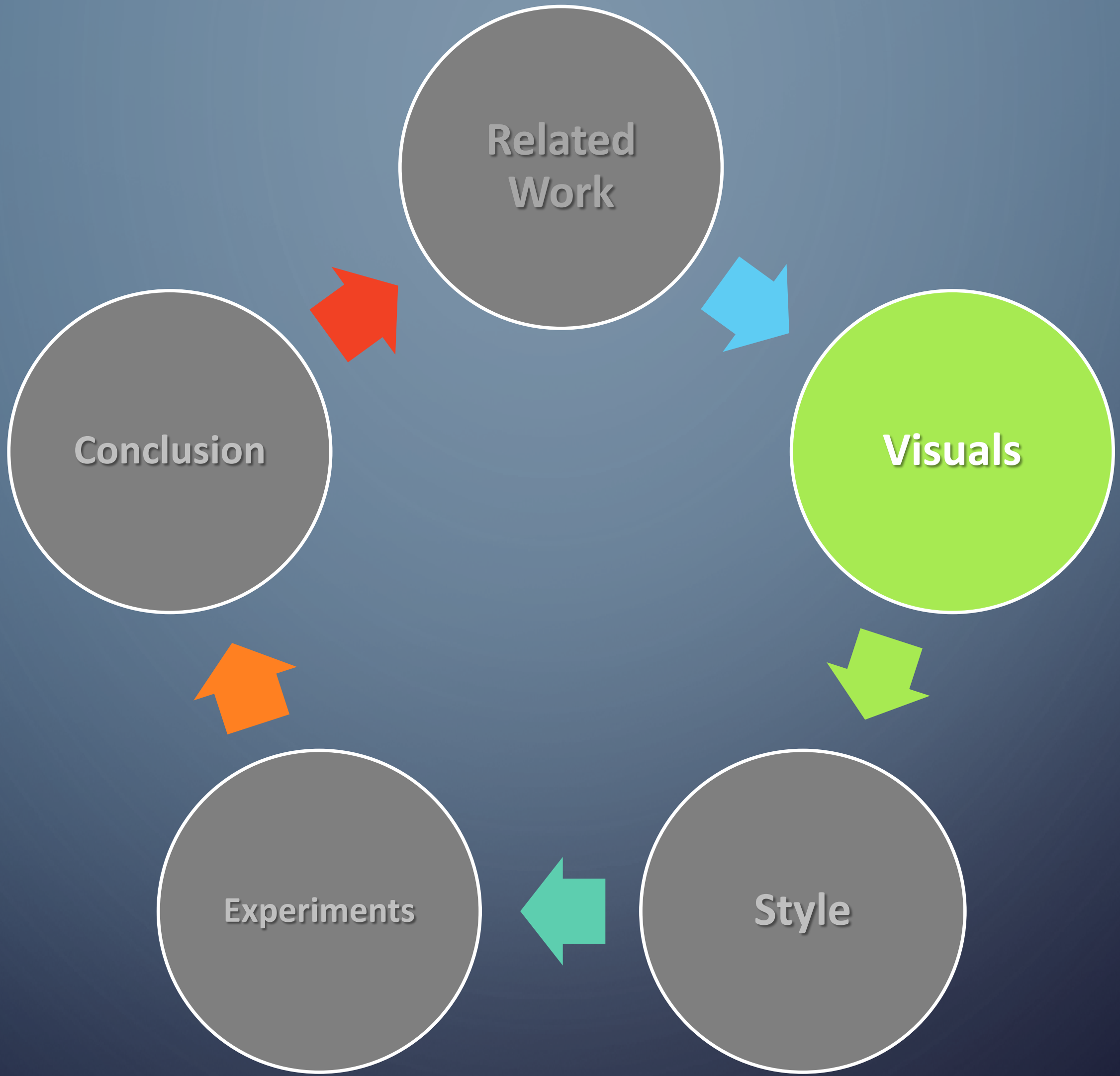
Table 3.1: A tentative classification of recent approaches on image deblurring.

Example taken from Xiaopeng Zhang [5]

IF NO SPACE FOR TABLE

Arguably, the closest existing method is Tensorfaces, which has been used for multimodal decomposition, classification, dimension reduction, and synthesis [7, 11, 12]. MMDA can therefore be considered an alternative method. But as we will show, MMDA enjoys a number of advantages
over Tensorfaces: it is easier to understand and implement because it is based on standard linear algebra, rather than multilinear algebra; it is more efficient to compute, and better for mode-invariant classification, dimension reduction, and synthesis. We demonstrate these advantages
by proving MMDA's theoretical properties, and by running extensive experiments using face images.

Example taken from Sim et al. [6]



Related Work
Visuals
Style
Experiments
Conclusion

A PICTURE

paints

a thousand

WORDS

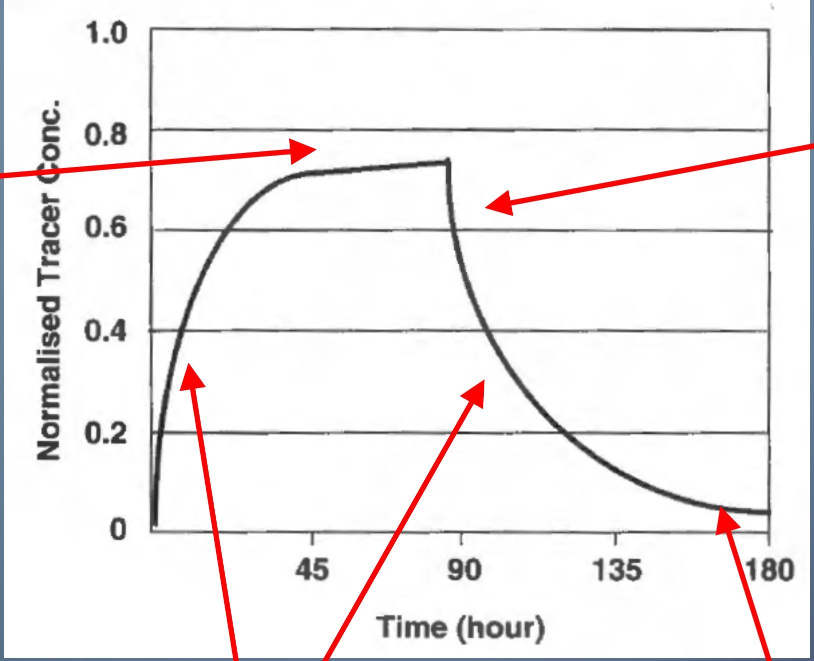
PRINCIPLES OF GOOD VISUALS

- A visual does not ask more questions than it can answer.
- A visual has its elements arranged to make its purpose immediately apparent.
- Besides the caption, a visual requires no external text support to be understood.

A VISUAL DOES NOT ASK MORE QUESTIONS THAN IT CAN ANSWER.

Figure 1. This is a red pen.

What questions does this figure ask?



Curve appears clipped here. Why?

What causes the sudden change here?

Curve is convex initially, but concave eventually. Explain.

Curve appears asymptotic. Why?

You need to answer these questions in the caption, otherwise reader will feel frustrated.

Example taken from Lebrun [2]

A VISUAL HAS ITS ELEMENTS ARRANGED TO MAKE ITS PURPOSE IMMEDIATELY APPARENT.

Methods	True-positive rate (%)	False-positive rate (%)
BN & BN	22.0	1.3
BN & MO	24.9	1.9
BN & MSV	39.2	0.2
PSY & BN	27.1	2.6
PSY & MO	27.0	2.7
PSY & MSV	66.9	0.3
COR & BN	23.0	1.9
COR & MO	25.8	2.5
COR & MSV	38.1	0.2
BN	21.8	1.2
MO	24.8	1.9
MSV	35.9	0.2

Example taken from Lebrun [2]

RE-ARRANGED, AND BOLDED

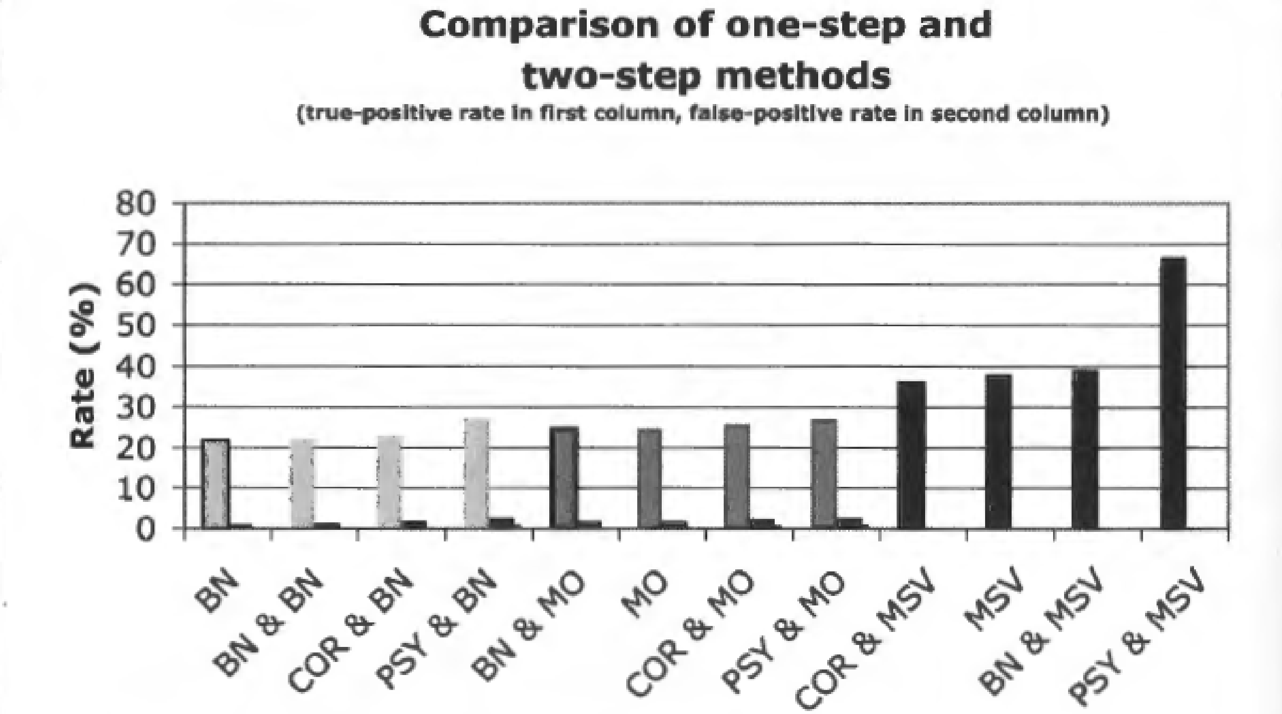
Methods (1 step & 2 steps)	True-positive rate (%)	False-positive rate (%)
BN	21.8	1.2
BN & BN	22.0	1.3
COR & BN	23.0	1.9
PSY & BN	27.1	2.6
MO	24.8	1.9
BN & MO	24.9	1.9
COR & MO	25.8	2.5
PSY & MO	27.0	2.7
MSV	**35.9**	**0.2**
COR & MSV	38.1	0.2
BN & MSV	39.2	0.2
PSY & MSV	**66.9**	**0.3**

Now, we can easily understand that

- Adding a 2nd step to BN, MO results in minor improvement only
- One step MSV method is superior to one-step BN, MO methods
- PSV + MSV almost doubles true-positive rate

Example taken from Lebrun [2]

SAME DATA, AS A BAR CHART



☛9. **Visual gallery of honours: the clear diagram.** ☛7 (modified). The comparison of one-step and two-step methods reveals three facts: (1) the improvement resulting from the addition of a second step to the BN and MO methods is minor; (2) the one-step MSV method (35.9% true-positive, 0.2% false-positive) is superior to the one-step BN and MO methods; and (3) adding the PSY method as a second step to MSV provides close to a twofold increase in performance (66.9% true-positive).

Example taken from Lebrun [2]

Example taken from Lebrun [2]

BESIDES THE CAPTION, A VISUAL REQUIRES NO EXTERNAL TEXT SUPPORT TO BE UNDERSTOOD.

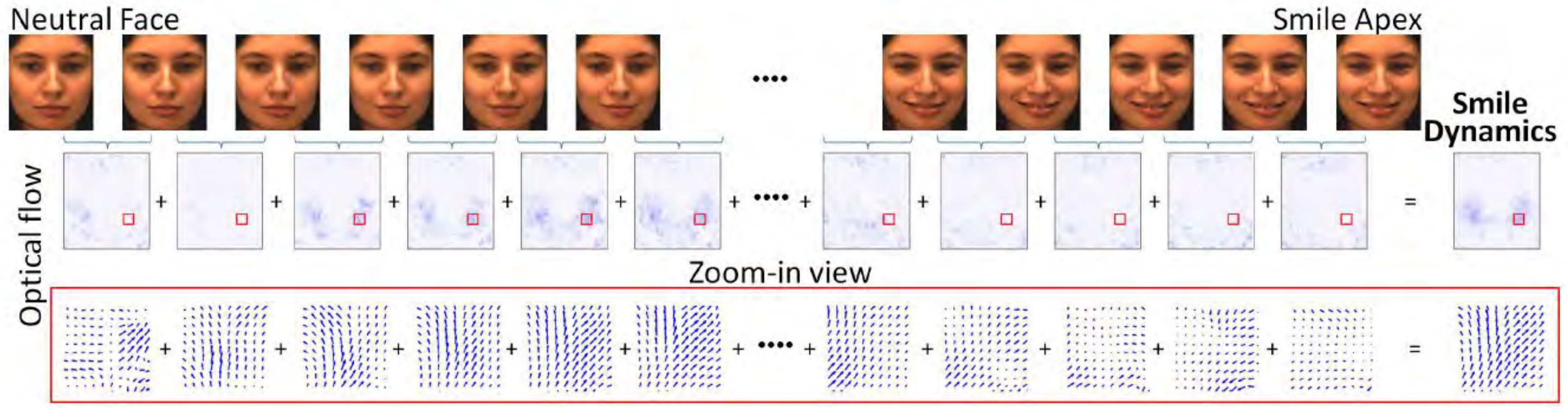
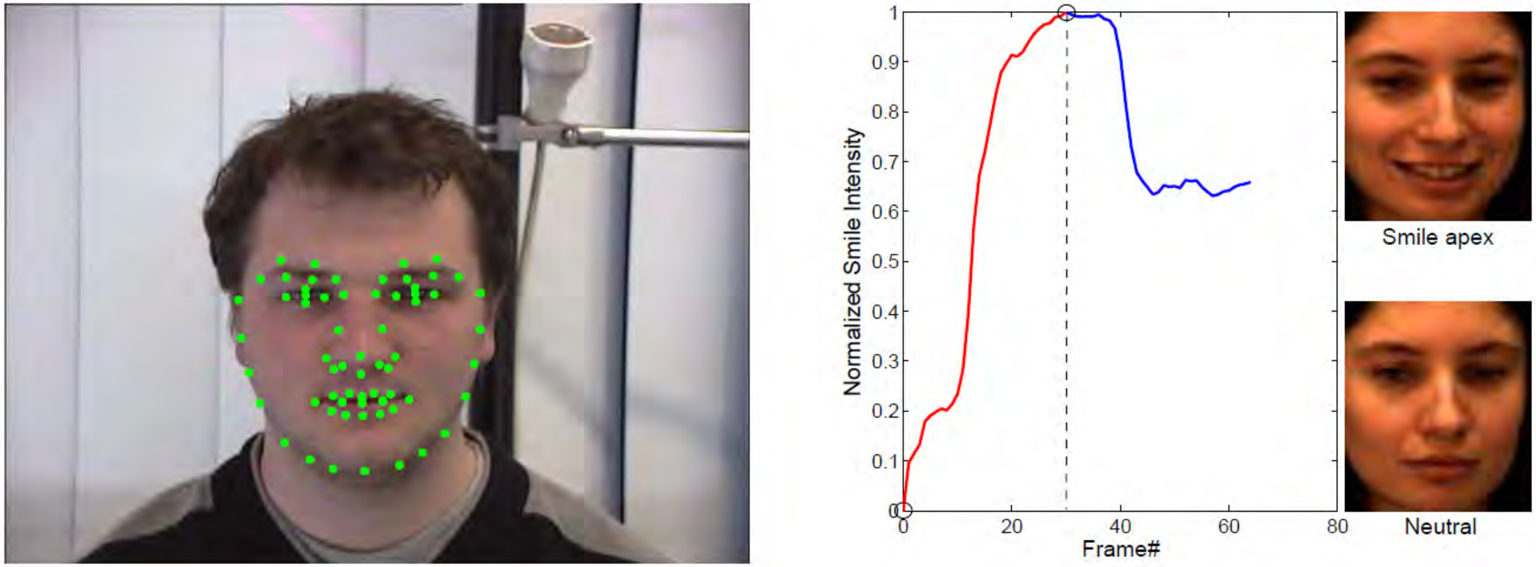


Figure 3.1: Smile dynamics is defined as the sum of a series of optical flow fields which are computed from the pairs of neighboring frames of a smile video.

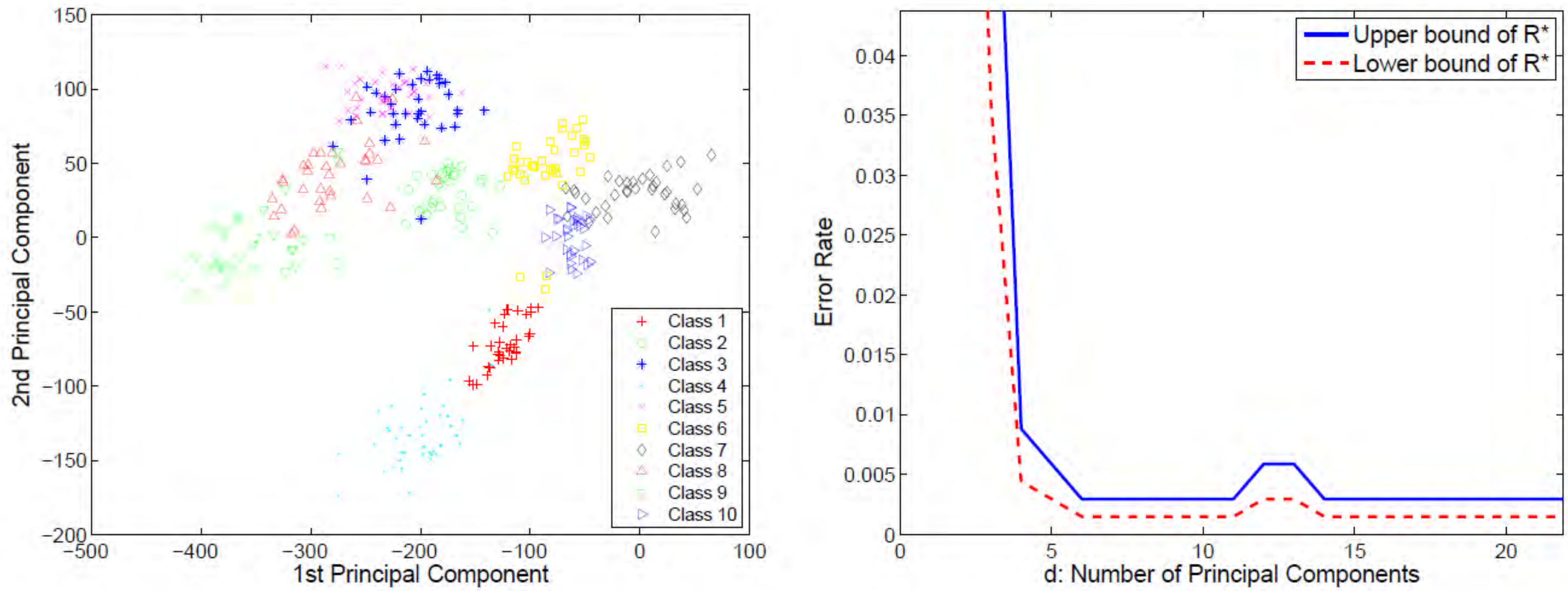
Example taken from Ning Ye [7]



(a) (b)

Figure 3.2: (a) Face localization result; (b) Normalized smile intensity: the red and the blue curves illustrate the neutral-to-smile period and the smile-to-neutral period, respectively; the neutral face and smile apex images are shown on the right.

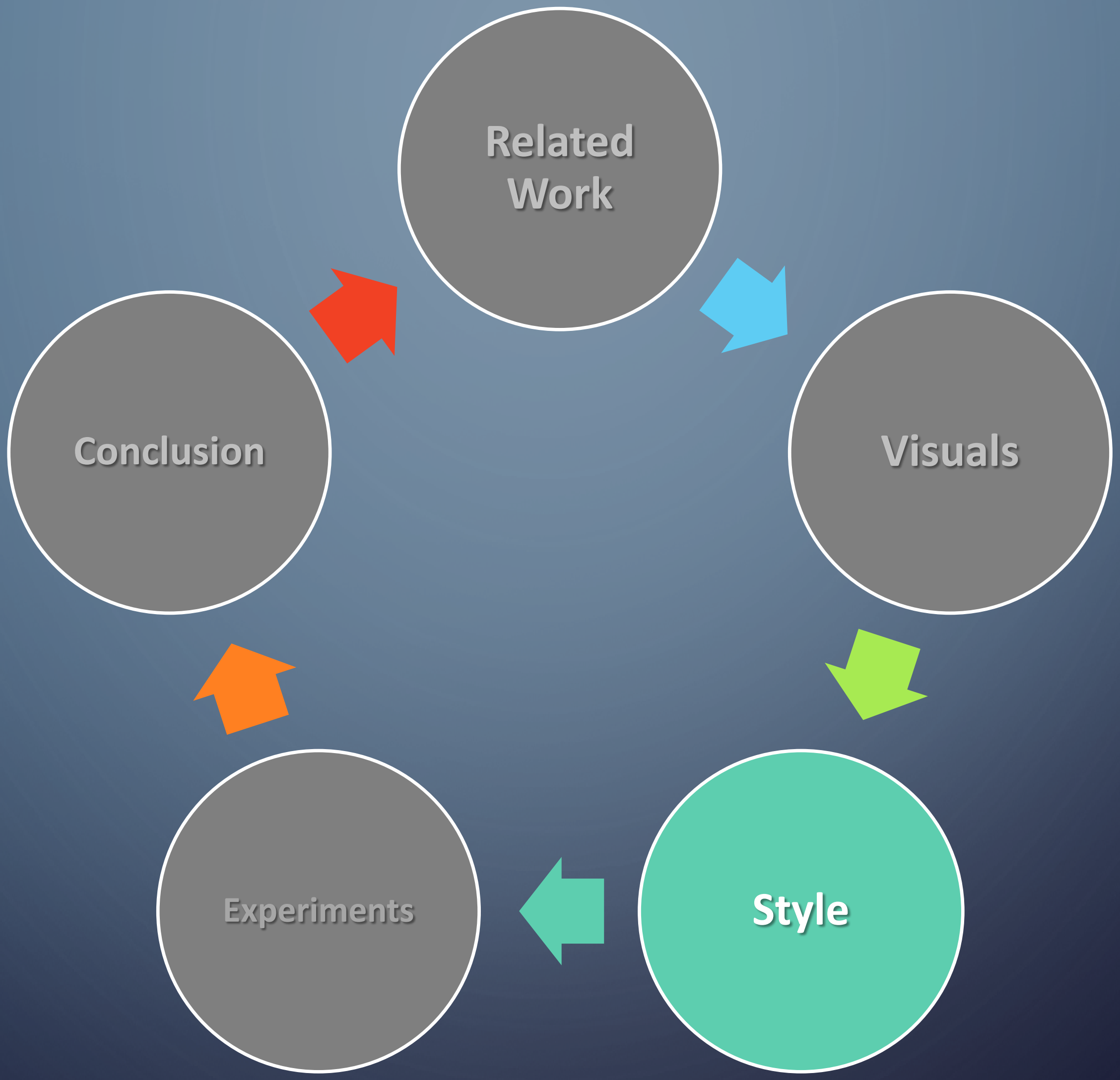
Example taken from Ning Ye [7]



(a) (b)

Figure 3.4: Class separability studies: (a) Data visualization after projected to 2D space; (b) The band of R^*: the Bayes' error rate R^* is bounded by the blue curve and the red dashed curve (Eq.(3.5));the horizontal axis denote the number of principal components d used in dimension reduction (Eq.(3.3)).

Example taken from Ning Ye [7]



Related Work
Visuals
Style
Experiments
Conclusion

WRITING LISTS

Given that the adversary have complete access to the original face database, he may conduct 3 types of attack:

1. If the task to be performed can be completed with reasonable accuracy by a computer algorithm, he can simply just do that. This is the most direct attack possible.

2. If the task is difficult for a computer, the adversary may however figure out the inverse function of each distorted image and obtain the original images.

3. The 3rd type of attack is to be deployed if the distortion is not reversible. By recruiting human solvers, he exhaustively matches every distorted image with the original.

Inconsistent phrasing!

Example taken from Chua et al. [8]

REVISED

There are 3 types of such attacks:

1. The black-box attack. This is where the adversary uses an algorithm to attack the face CAPTCHA, treating it like a black-box.
2. The distortion reversal attack. This is where the adversary attacks the face CAPTCHA by using an algorithm to obtain from the distorted images, images close to the original.
3. The human solver attack. This is where the adversary recruits human solvers to exhaust all images appearing the face CAPTCHA ahead of time.

Consistent phrasing.
Achieved by defining names.

Example taken from Chua et al. [8]

USING REPEATED STRUCTURE TO AID COMPARISON

There are 3 types of such attacks:

1. The black-box attack. This is where the adversary uses an algorithm to attack the face CAPTCHA, treating it like a black-box. The attack is used when the CAPTCHA can be automatically solved without requiring any knowledge of the kinds of distortions the original images undergo. Based on the algorithms used, …. For instance, while face detection algorithms may not be able to detect the face, …

Explanation

Name

Example

Usage

Example taken from Chua et al. [8]

USING REPEATED STRUCTURE TO AID COMPARISON

Explanation

2. The distortion reversal attack. This is where the adversary attacks the face CAPTCHA by using an algorithm to obtain, from the distorted images, images close to the original. … This is useful when the CAPTCHA task can be easily performed by existing algorithms on the original images. For instance, the task of …

Name

Example

Usage

Example taken from Chua et al. [8]

USING REPEATED STRUCTURE TO AID COMPARISON

Explanation

3. The human solver attack. This is where the adversary recruits human solvers to exhaust all images appearing the face CAPTCHA ahead of time. … This attack requires tedious pre-processing and is usually used when the above 2 attacks fail. … For instance, Microsoft uses a cat/dog image …

Name

Example

Usage

Example taken from Chua et al. [8]

AVOIDING SYNONYMS

Synonyms confuse your reader.
Use words consistently.

AVOIDING SYNONYMS

Multi-modal biometrics are shown to perform better than uni-modal biometrics by using ***fusion techniques*** at different *layers* (9). The most common ***methods of fusion*** as described in (9) are -

- Feature *level*: combining the feature vectors of different modalities to learn a single model of the user. For example, concatenate face, iris, voice feature vectors to a single classifier.
- Score *level*: combining the scores of different classifiers, where typically there is at least one classifier for one modality. For example, averaging/weighted averaging of scores from the classifiers and matching against a threshold to decide.
- Decision *level*: - combining the decisions of multiple classifiers, and using techniques like majority voting.

techniques = methods?   layers = level?

Example taken from Divya Sivasankaran [9]

EXCEPT PRONOUNS

Pronouns (eg. this, it, their) are acceptable as synonyms.

But be careful!

The cellular automaton (CA) cell, a natural candidate to model the electrical activity of a cell, is an ideal component to use in the simulation of *intercellular communications*, such as those occurring between cardiac cells, and to model *abnormal asynchronous propagations*, such as *ectopic beats*, initiated and propagated cell-to-cell, regardless of the complexity of *their* patterns.

What does *their* refer to?

Example taken from Lebrun [2]

OMIT PRONOUNS TO CLARIFY

The cellular automaton (CA) cell -- a natural candidate to model the electrical activity of a cell -- is an ideal component to use in the simulation of intercellular communications, such as those occurring between cardiac cells, and to model *the cell-to-cell initiation and propagation of abnormal asynchronous events (such as ectopic beats) with or without complex patterns*.

Example taken from Lebrun [2]

REPEAT NOUNS TO CLARIFY

The cellular automaton (CA) cell, a natural candidate to model the electrical activity of a cell, is an ideal component to use in the simulation of intercellular communications, such as those occurring between cardiac cells, and to model *abnormal asynchronous events, such as ectopic beats, initiated and propagated cell-to-cell, however complex the propagation pattern may be.*

Example taken from Lebrun [2]

MINIMIZE USAGE OF PASSIVE VOICE

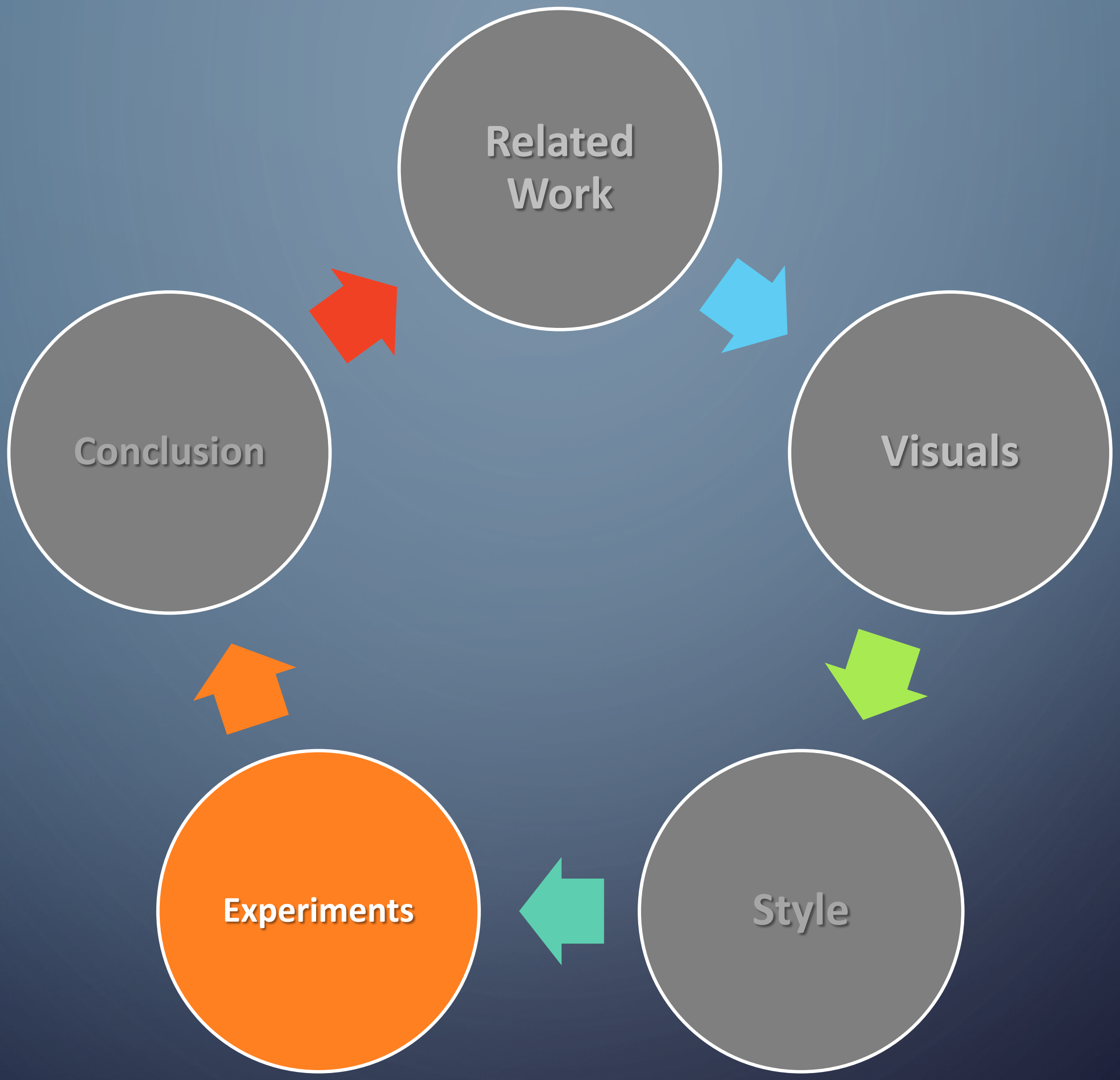
The fingerprint and face data were captured and processed by image pre-treatment.

Was the pre-treatment done by
(a) other researchers previously,
(b) the author in a prior paper, or
(c) the author in the current paper?

The passive voice cannot answer this question.

PREFER THE ACTIVE VOICE

We pre-treat fingerprint and face data using the method in our previous paper (10).


Related Work
Visuals
Style
Experiments
Conclusion

SET EXPERIMENTAL GOALS, THEN FULFILL THEM

VI. Experiments & Discussion

How good is our face alteration? More precisely,

Q1. When we alter a facial attribute, say, gender, is it effective?

Q2. When we alter one facial attribute but retain others, are the unchanged attributes perceived as such?

Q3. Does increasing the intensity of a parameter manifest in a corresponding increase in the attribute?

Q4. When we alter identity, is it effective?

To answer these questions, we will use a Change Detector (CD), *i.e.* a vision algorithm, to compare an original face image with its altered image. This is in line with our motivation to protect privacy while allowing visual analytics (*i.e.* other computer vision algorithms) to function normally. In all our experiments, we use a set of test images that are different from our training set.

A. *Evaluation metric*

We build several CDs, one each for identity, gender, race and age. Each CD accepts two inputs, an original face

Example taken from Sim and Zhang [11]

Example taken from Sim and Zhang [11]

SET EXPERIMENTAL GOALS, THEN FULFILL THEM

B. Experiments on single attribute change

To answer questions Q1 and Q2, we changed one facial attribute while retaining the other two. We generated between

C. Experiments on multiple attribute change

We next examine the effect of changing two or more facial attributes. Table II summarizes the β values. Again, the

D. Experiments on identity change

In fact, identity change can be easily observed in our experimental results. To validate this, we asked 5 volunteers to compare the identities in an original image and its altered

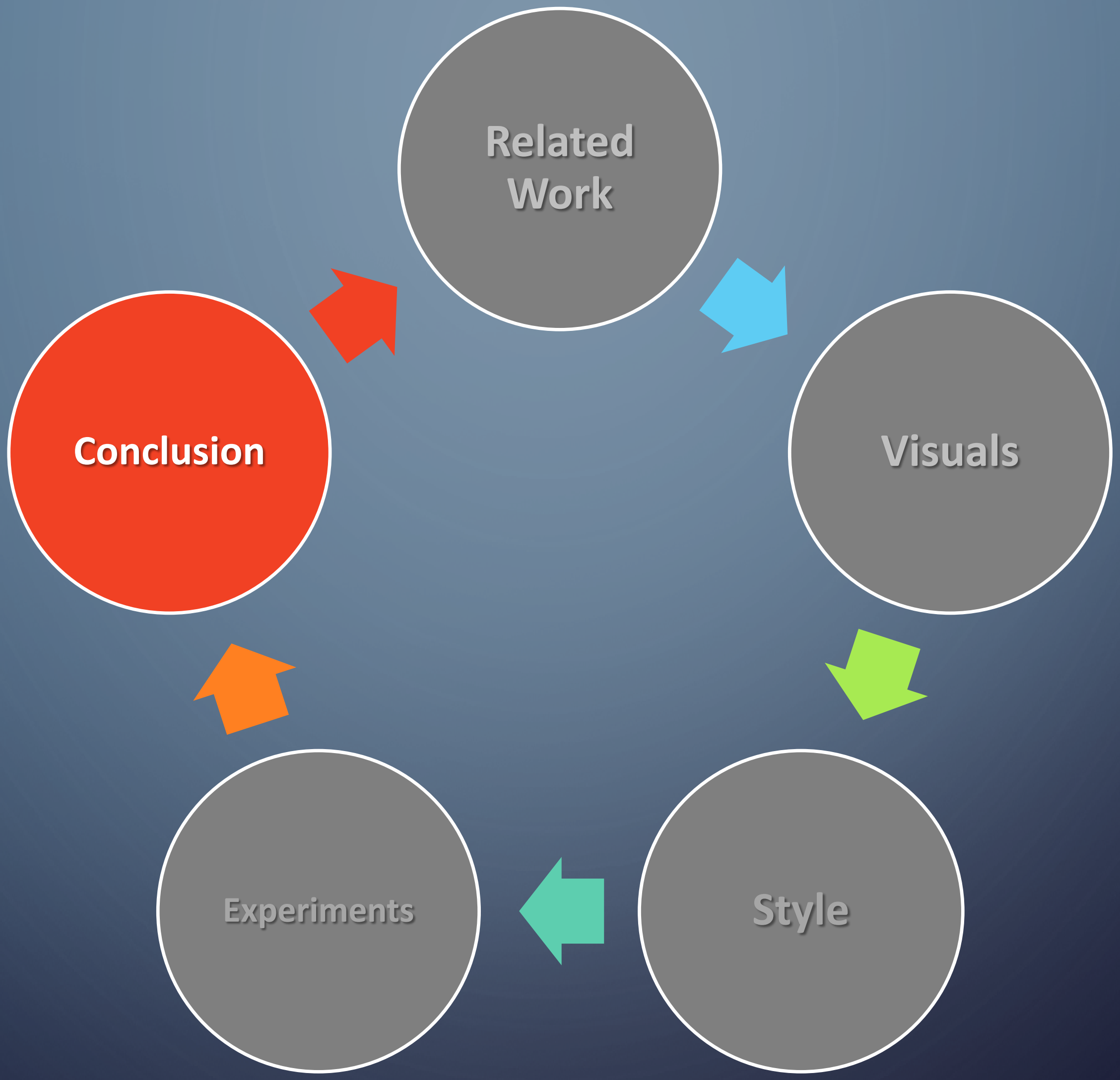
Example taken from Sim and Zhang [11]

SET EXPERIMENTAL GOALS, THEN FULFILL THEM

E. Discussion

1. From all these experiments, we conclude that our method is effective in altering the facial attributes of gender, race, age, and identity, whether singly or in different combinations. Question Q4 is answered in the affirmative by Table III; while Q1, Q2 and Q3 are all answered in the affirmative by Tables I and II.
2. We could not compare with existing works because ours is the first to selectively alter some facial attributes while retaining other attributes. There is no prior work to compare to.

Example taken from Sim and Zhang [11]


Related Work
Visuals
Style
Experiments
Conclusion

PURPOSE OF THE CONCLUSION

1. It restates the contribution, with a particular emphasis on what it allows others to do.
2. It proposes new research directions to prevent duplication of effort or to encourage collaboration.

EXAMPLE

To the best of our knowledge, our template protection scheme is the first of its kind. This ability to guarantee irreversibility, revocability, and unlinkability for any Face Verifier, while maintaining good verification performance, has not been reported in the literature. We achieve this by rendering user specific virtual faces, which are carefully placed far apart from one another in MMDA's identity subspace. While our experimental results on OpenBR and OpenFace are encouraging, it would be nice to provably guarantee that performance will not worsen for all Face Verifiers. We intend to pursue this in future work. Another area of improvement is to remove the capacity limit (see Section III-C) in our scheme, so that infinitely many revocations are permitted. Still another improvement is to guarantee irreversibility when both the token and virtual face are stolen.

Example taken from Li et al. [10]

C.F. INTRODUCTION

Our template protection scheme makes two contributions:

(a) it possesses the properties of irreversibility, revocability, unlinkability and good verification performance, in the case of non-pairwise exposure of token and virtual biometric data;

(b) it may be added on to any Face Verifier, because it treats the Verifier as a black box, requiring only that the Verifier outputs a score between 0 and some maximum value τ .

Example taken from Li et al. [10]

WORK IN SOME EMOTIONS

We are *pleased* to present the novel concept of Controllable Face Privacy for the nuanced protection of face images. Applying multimodal discriminating analysis on our face encoding scheme results in a Semantic basis with which we may decompose a face into its gender, race and age attributes. In turn, this permits the synthesis of novel faces with new, desired attributes. Moreover, privacy protection mechanisms, such as k-anonymity, L-diversity, t-closeness, are easily incorporated into our method, thereby providing provable guarantees on our altered faces. In the near future, we intend to get human volunteers to assess the quality of our altered images.

Example taken from Sim and Zhang [11]

C.F. INTRODUCTION

Our contribution: This paper pioneers the notion of Controllable Face Privacy for the protection of privacy in face images. The key idea is to selectively alter some facial Attributes while retaining others. To this end we employ a subspace decomposition technique to decouple the parameters that control different facial attributes. In each subspace, we may then independently vary the said parameters and then synthesize faces with new attributes. This not only permits the privacy protection of facial identity (which is the sole concern of all existing work), but also of gender, race and age as well. Furthermore, we show that we can easily incorporate the mechanisms of k-anonymity, L-diversity, and t-closeness [13] (pioneered by the data mining research community) to provide provable privacy guarantees on the altered faces. We run extensive experiments — we tested our altered images on Face++, a commercial face analysis software that can classify gender, age and race — to show that our alteration is indeed effective.

Example taken from Sim and Zhang [11]

BRING IN THE HUMAN INTEREST STORY, IF POSSIBLE

I. Introduction

In 1984, based on the testimony of five eyewitnesses, Kirk Bloodsworth was convicted of the rape and murder of a nine-year-old girl and sentenced to the gas chamber. After Bloodsworth served nine years in prison, DNA testing proved him to be innocent [1]. Such devastating mistakes by eyewitnesses are not rare, and more than 75% of the convictions overturned through DNA testing since the 1990s were based on eyewitness testimony [2].

The eyewitness testimony for forensic applications has had a long history, with roots that go back to the beginning

Example taken from Nejati and Sim [12]

BRING IN THE HUMAN INTEREST STORY, IF POSSIBLE

V. Summary and Conclusion

More than 30 years of psychological studies show that forensic sketches are highly unreliable due to problems such as verbal overshadowing in the very first steps of eyewitness testimony methods (ETMs) [7], and piecewise reconstruction in the next steps [10]. However, no practical

The main motivation for this work was to create the missing link between psychological findings, automatic face sketch recognition, and real world applications, and therefore reduce the chance of wrongful convictions of innocents, such as the Kirk Bloodsworth. We hope that this work serves as a first step for better methods benefiting both computer vision and forensic sciences.

Example taken from Nejati and Sim [12]

FINALLY …

GOOD WRITING IS HARD WORK!

REFERENCES

1. Jeffrey McQuain. “Power Language: Getting the Most out of Your Words.” Houghton Miffin 1996.
2. Jean-Luc Lebrun. Scientific Writing: A Reader and Writer’s Guide. World Scientific, 2011.
3. Janakiraman, Rajkumar, and Terence Sim. "Keystroke dynamics in a general setting." Advances in Biometrics (2007): 584-593.
4. Xiaopeng Zhang. “Gradient Variation: a Key to Enhancing Photographs Across Illumination”. Ph.D. Dissertation, School of Computing, National University of Singapore, 2009.
5. Xiaopeng Zhang. “Deblurring in Digital Photography”. Graduate Research Paper, School of Computing, National University of Singapore, 2005.
6. Sim, Terence, Sheng Zhang, Jianran Li, and Yan Chen. "Simultaneous and orthogonal decomposition of data using multimodal discriminant analysis." In Computer Vision, 2009 IEEE 12th International Conference on, pp. 452-459. IEEE, 2009.
7. Ning Ye. “Personal Identification from Facial Expression.” Ph.D. Dissertation, School of Computing, National University of Singapore, 2010.
8. James Chua, Hossein Nejati, and Terence Sim. “Face CAPTCHA: Design Guidelines and Implementatons”. Unpublished. 2010.
9. Divya Sivasankaran. “Context Invariant Fusion for Multi-modal Biometric Authentication.” CS6240 Project Proposal, School of Computing, National University of Singapore, 2017.
10. Jing Li, Yongkang Wong, and Terence Sim. “Towards Protecting Biometric Templates Without Sacrificing Performance.” International Conference on Pattern Recognition, 2016.
11. Sim, Terence, and Li Zhang. "Controllable face privacy." In Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on, vol. 4, pp. 1-8. IEEE, 2015.
12. Nejati, Hossein, Terence Sim, and Elisa Martinez-Marroquin. "Do you see what i see? A more realistic eyewitness sketch recognition." In Biometrics (IJCB), 2011 International Joint Conference on, pp. 1-8. IEEE, 2011.