質感のつどい 2015/11/25

質感研究の現在

西田眞也

ＮＴＴコミュニケーション科学基礎研究所

質感認識

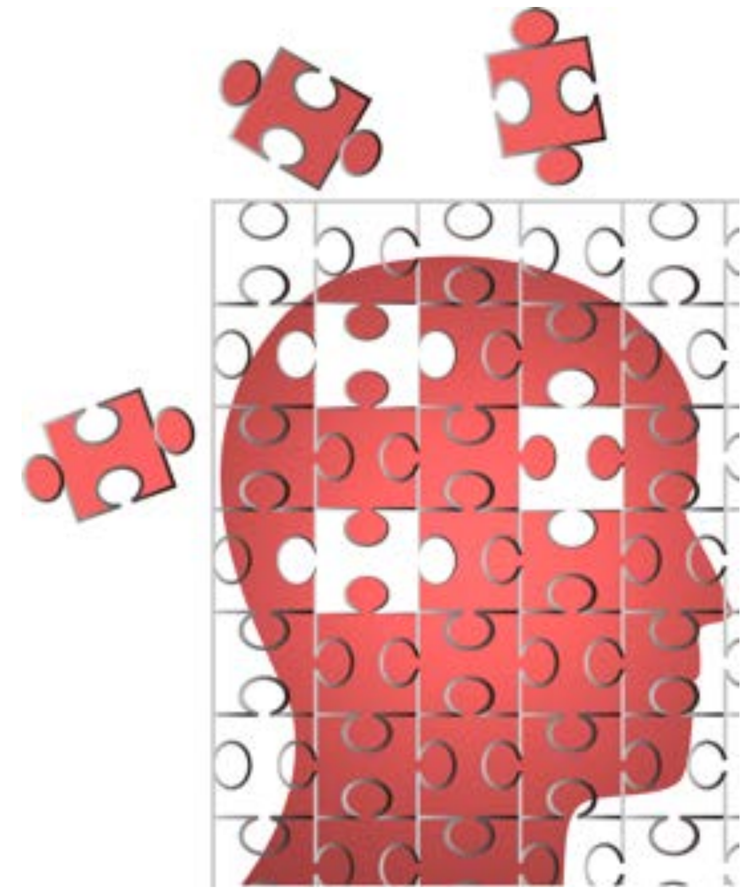
- 五感を通した脳による物体の本性の解読
  - **物性**（例：光沢感、透明感）
  - **材質**（例：陶器、金属）
  - **状態**（例：乾燥、凍結）
  - **感性的価値**（例：美味しそう）

質感研究の重要性

- **脳科学研究の大きなミッシングピース**

質感情報処理

対象認識 → 価値評価・行動選択 → 身体運動制御



- **情報工学の技術革新の宝庫**
  - ユーザが満足する質感をもったプロダクト
  - 質感を判断して適切に働くロボット

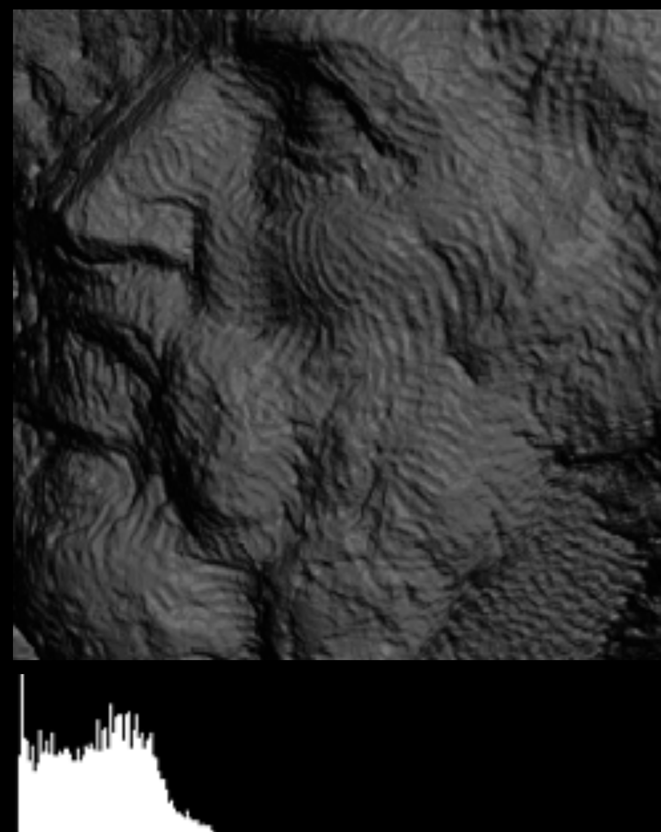
- **難問**
  - どういう**刺激入力**からどういう**質感反応**をどういう**計算原理**で推定？

疑問

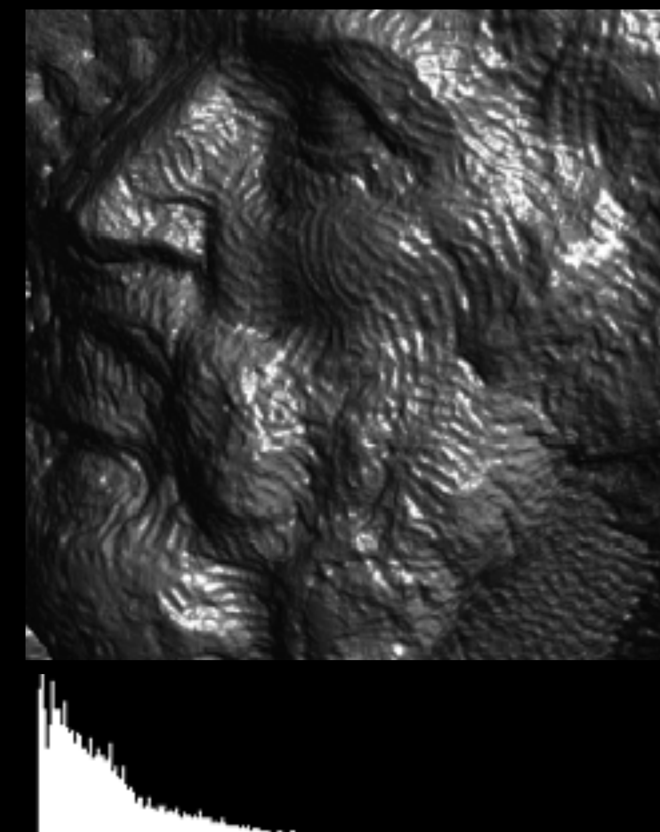
- どうして、物理的な光学シミュレーションを精緻にすると、本物の写真のようにリアルに見えるのか（質感が正しく再現されるのか）？
  - 脳が逆光学計算に基づいて画像生成の物理的正しさをチェックすることができるから？
  - 脳がリアリティを感じるのに必要な「画像特徴」が画像中にうまく再現されるようになるから？
- 脳の質感情報処理はどれほど深いのか？

光沢感の手がかり：画像統計量

平均輝度は同じ



輝度ヒストグラム



光沢のない明るい表面

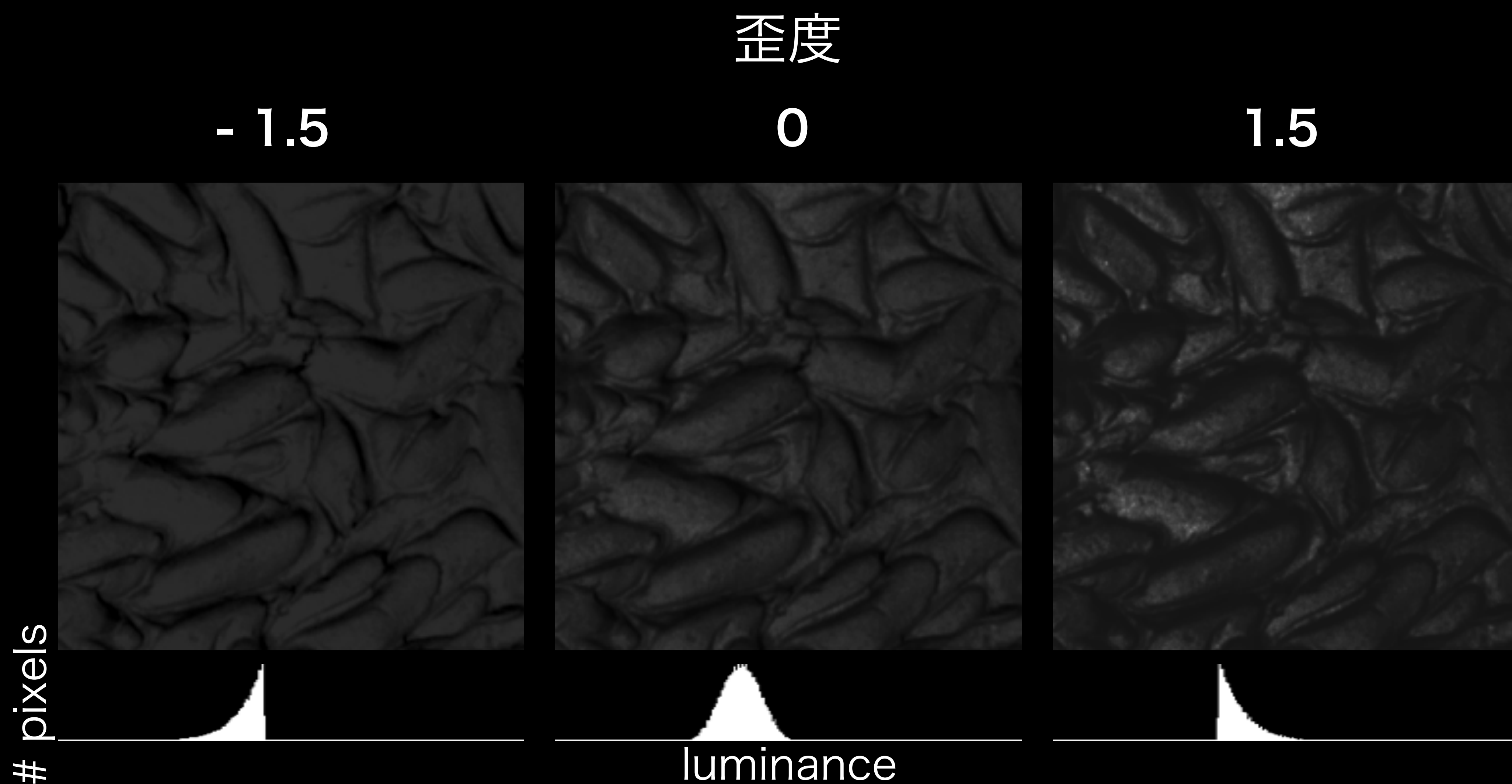
輝度ヒストグラムが
負の方向に歪む

光沢のある暗い表面

輝度ヒストグラムが
正の方向に歪む

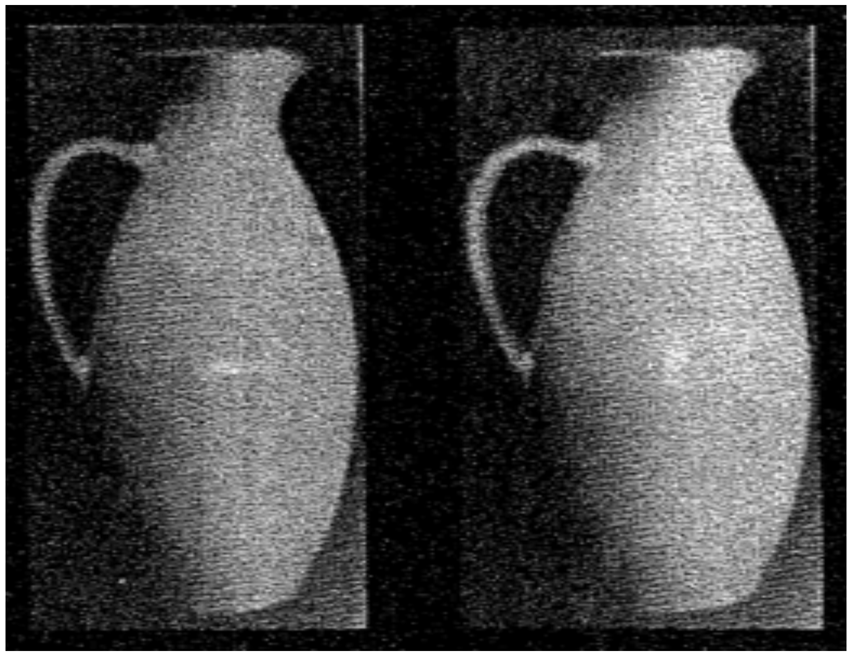
Motoyoshi, Nishida, Sharan, Adelson (2007), Nature, 447, 206-209

輝度ヒストグラムによる光沢感の操作

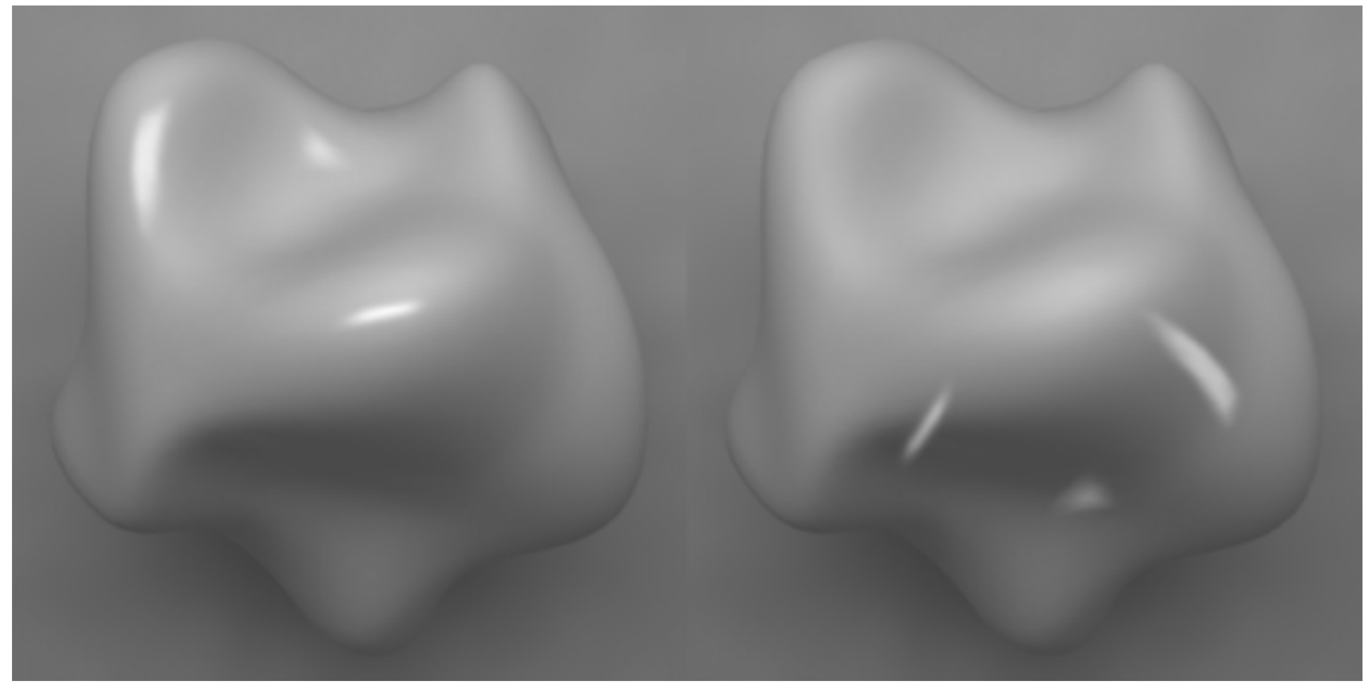


輝度ヒストグラムを人工的にゆがめると,
見かけの光沢や明るさが変化する

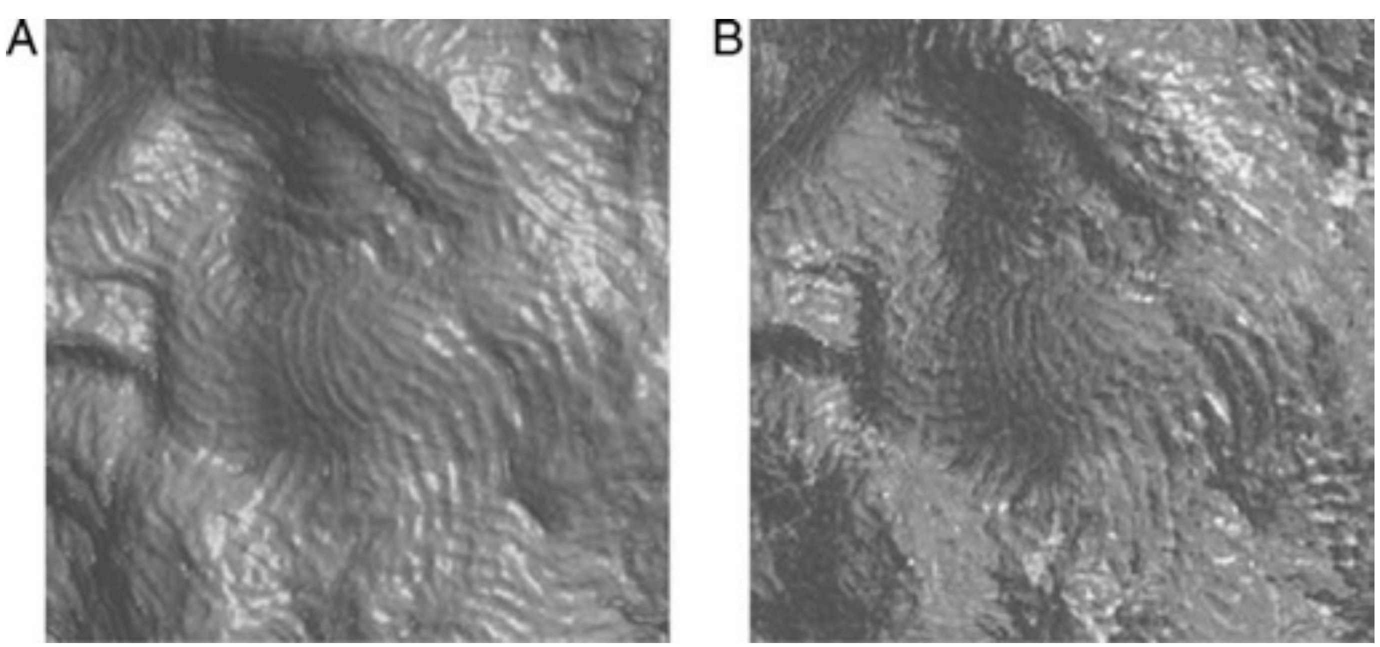
光沢感には形状情報も重要



Beck and Prazdny. Highlights and the perception of glossiness. Perception & psychophysics (1981) vol. 30 (4) pp. 407

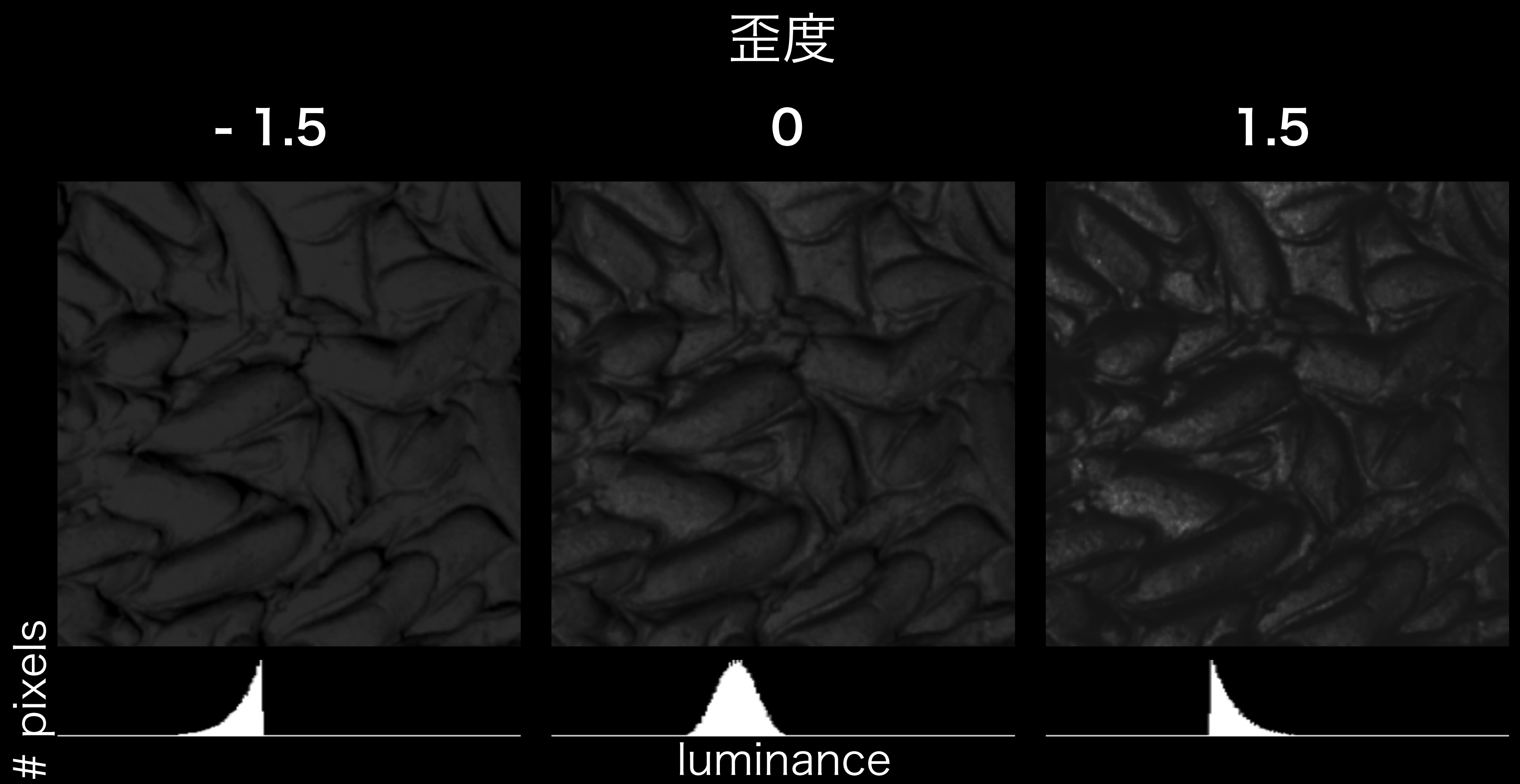


Todd, Norman & Mingolla, Lightness constancy in the presence of specular highlights. Psychological Science (2004) vol. 15 (1) pp. 33



Anderson, B., & Kim, J. (2009). Image statistics do not explain the perception of gloss and lightness. *Journal of vision*.

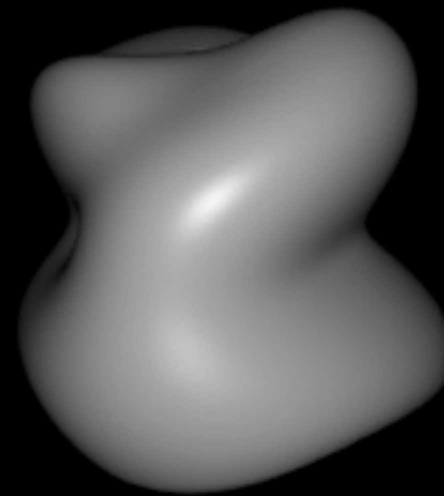
輝度ヒストグラムによる光沢感の操作

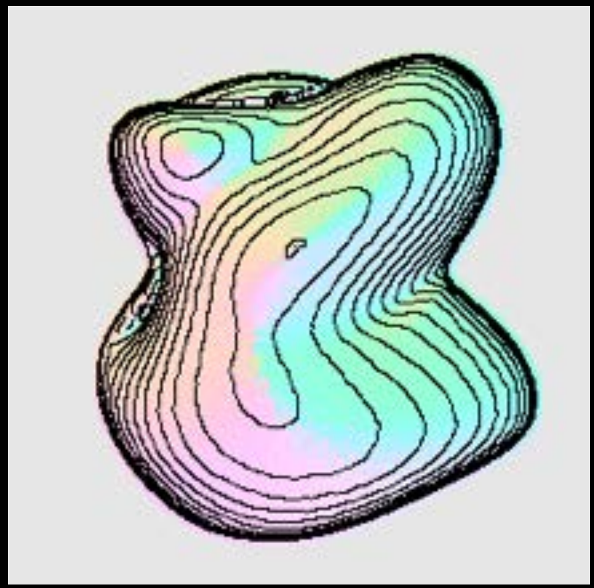


輝度ヒストグラムをゆがめても、
形状知覚は変化しない

輝度順序マップ

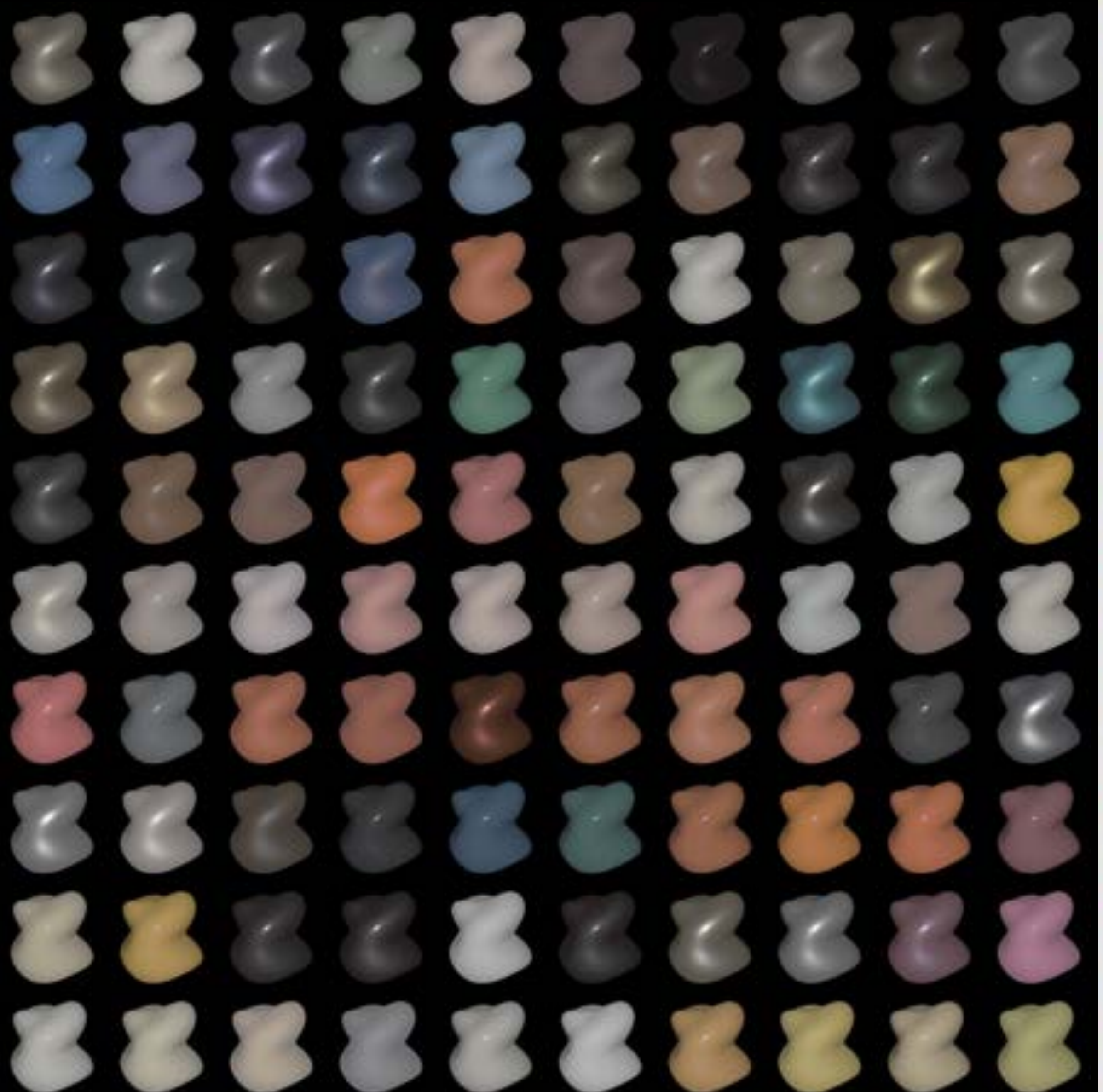
- 輝度ヒストグラムは輝度順序マップに影響しない
- 表面反射特性が変化しても輝度順序マップはほとんど変化しない



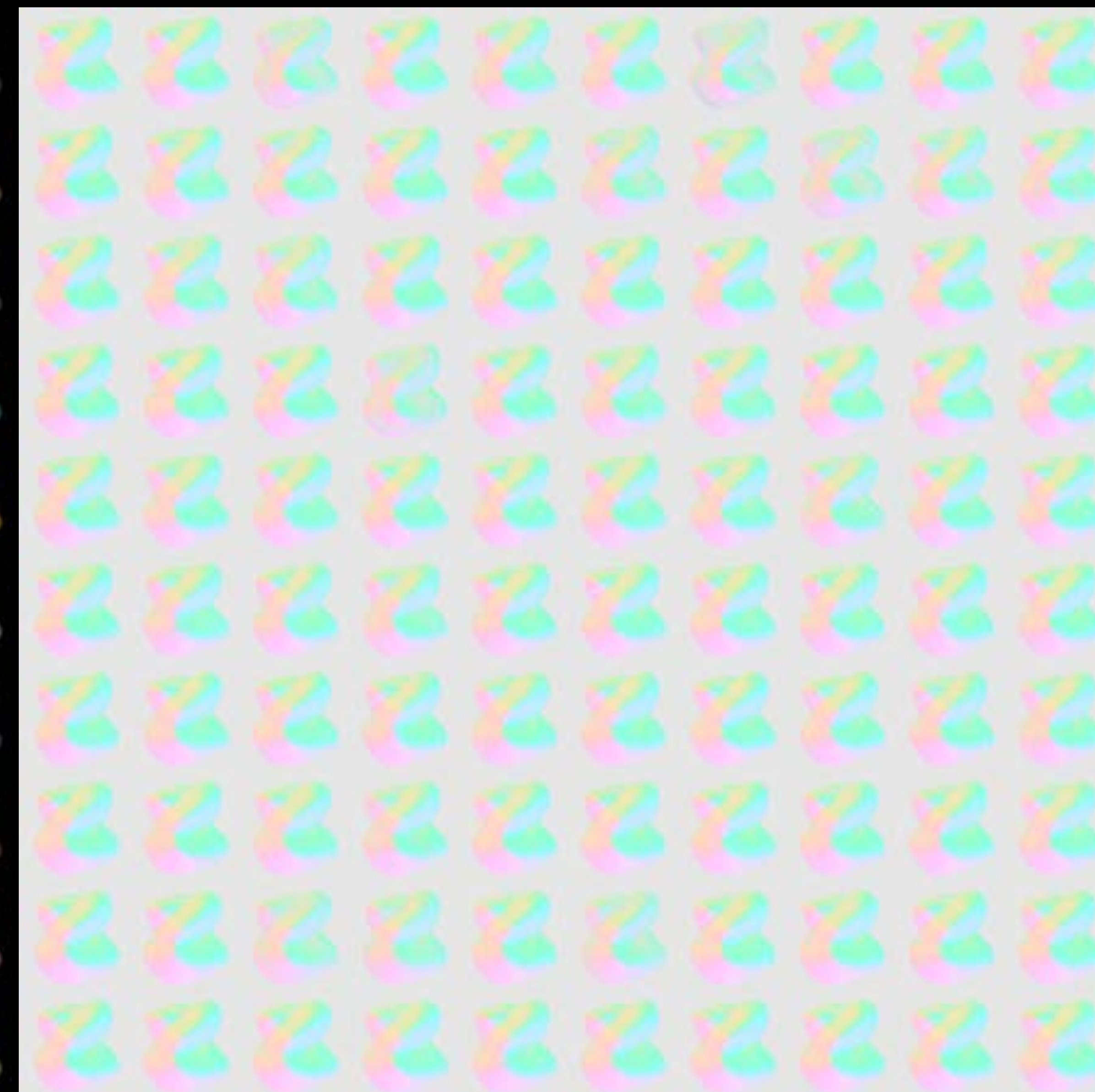


Intensity order is altered little by material

MERL BRDF



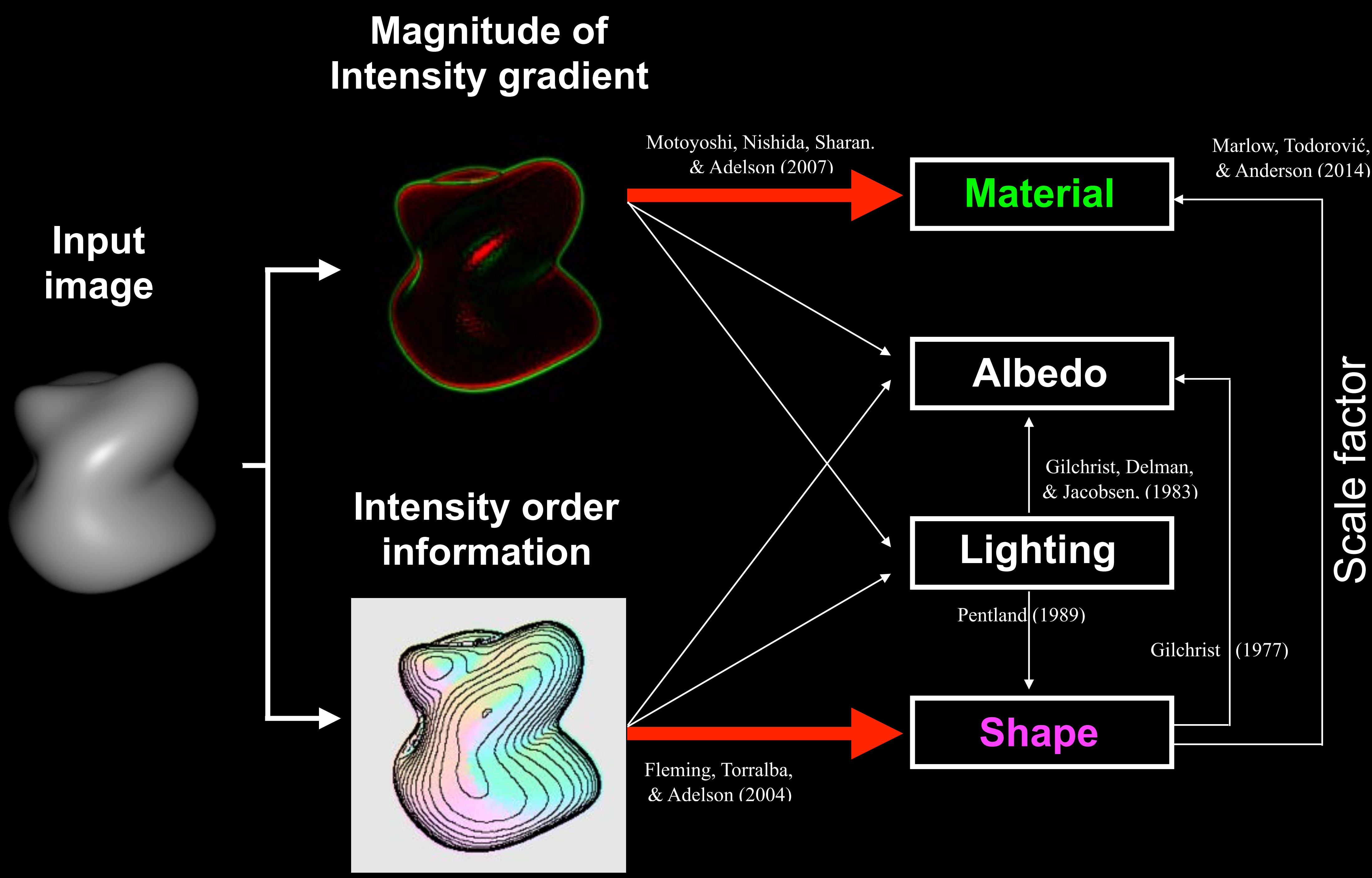
Intensity order information



Normalized intensity gradient map

Many material changes (in MERL database) have only minor effects on pixel intensity order.

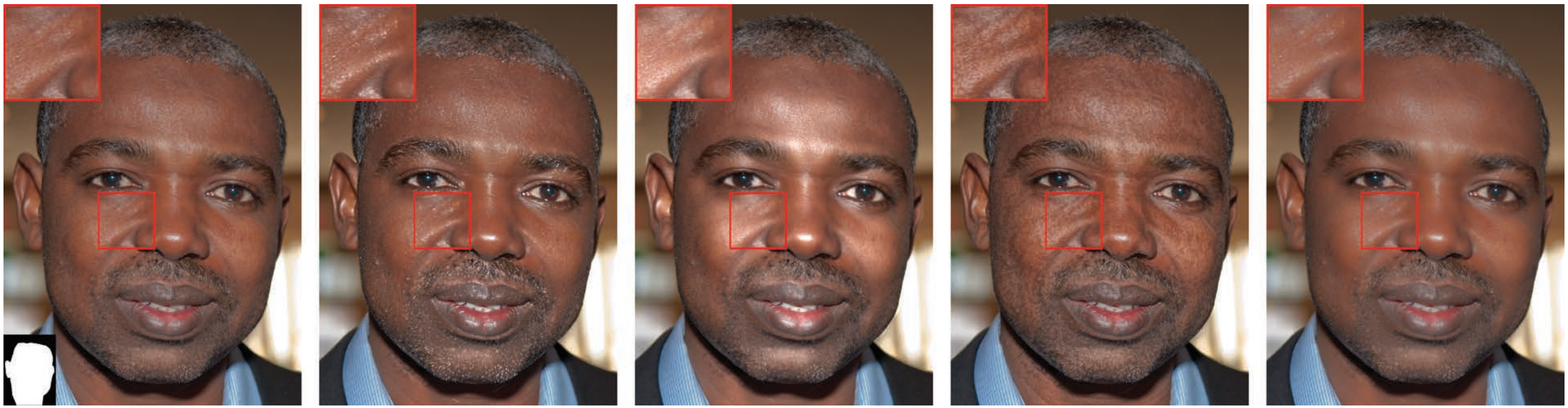
Separate processing of material and shape



(Sawayama & Nishida, *VSS 2014; APCV 2014; Nishida et al., VSS 2015*)

Image-based material editing

Boyadzhiev, I., Bala, K., Paris, S., & Adelson, E. (2015). Band-Sifting Decomposition for Image-Based Material Editing. ACM Transactions on Graphics, 34(5), 1–16.



(a) original image (b) wet/oily skin (c) smooth/shiny glow (d) more blemishes (old) (e) fewer blemishes (young)

Fig. 2. Our band-sifting operators are particularly useful for manipulating material properties in human faces. (a) Original image courtsey of Bychkovsky et al. [2011] with detail inset at upper left, and mask inset at lower left. (b) We sift and then boost the high-amplitude, positive coefficients in the high-spatial frequencies, which gives the skin a more shiny or wet look. (c) We manipulate the positive low-spatial frequencies coefficients, which gives the skin a soft glow. (d) We produce an aging effect by emphasizing blemishes and pores that are not noticeable in the input image. We achieve this by sifting and then boosting the low-amplitude coefficients in the high-spatial frequencies. (d) We reverse the effect, that is, reduce blemishes and pores, by decreasing the sifted coefficients from (d).

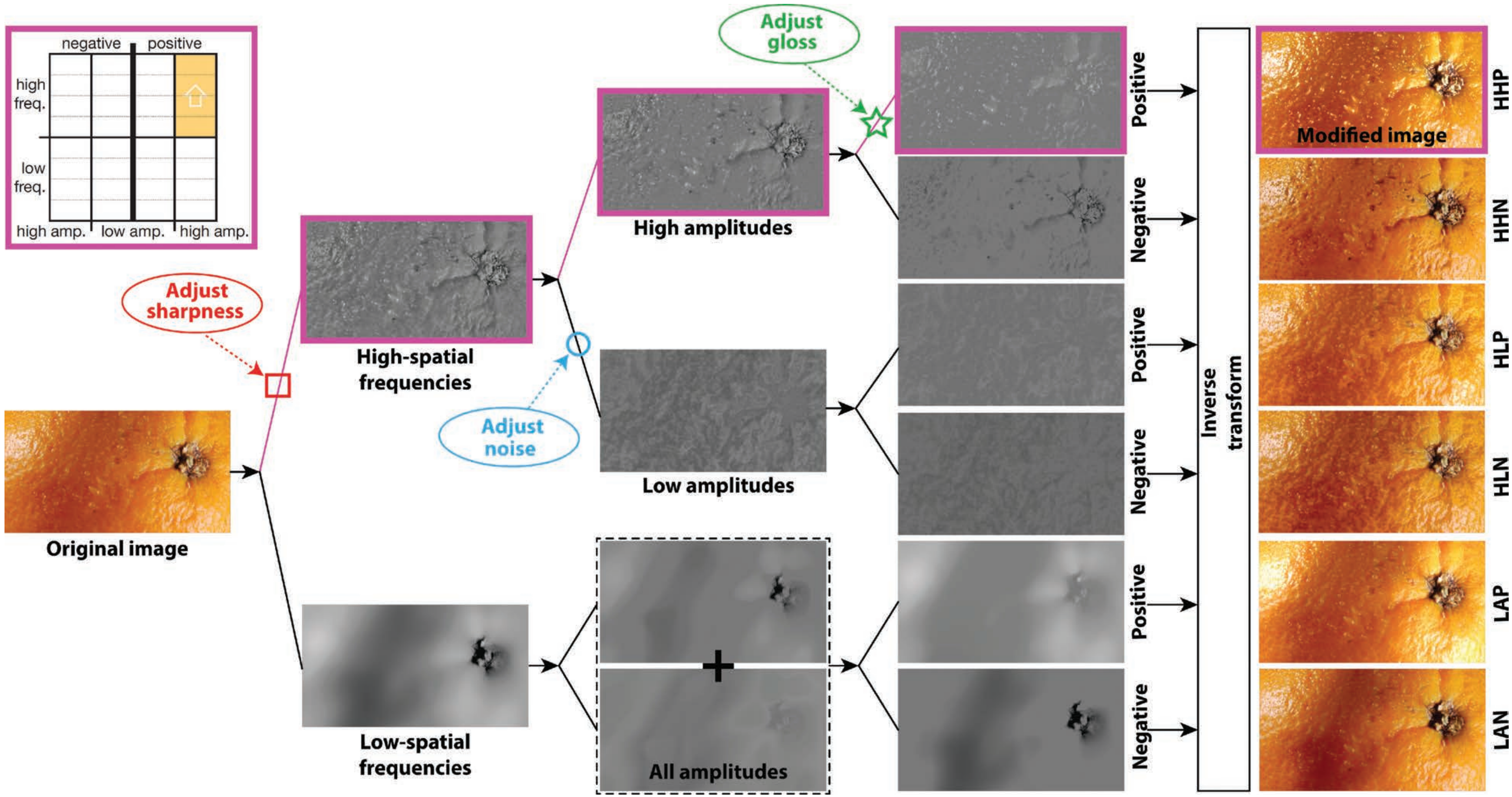
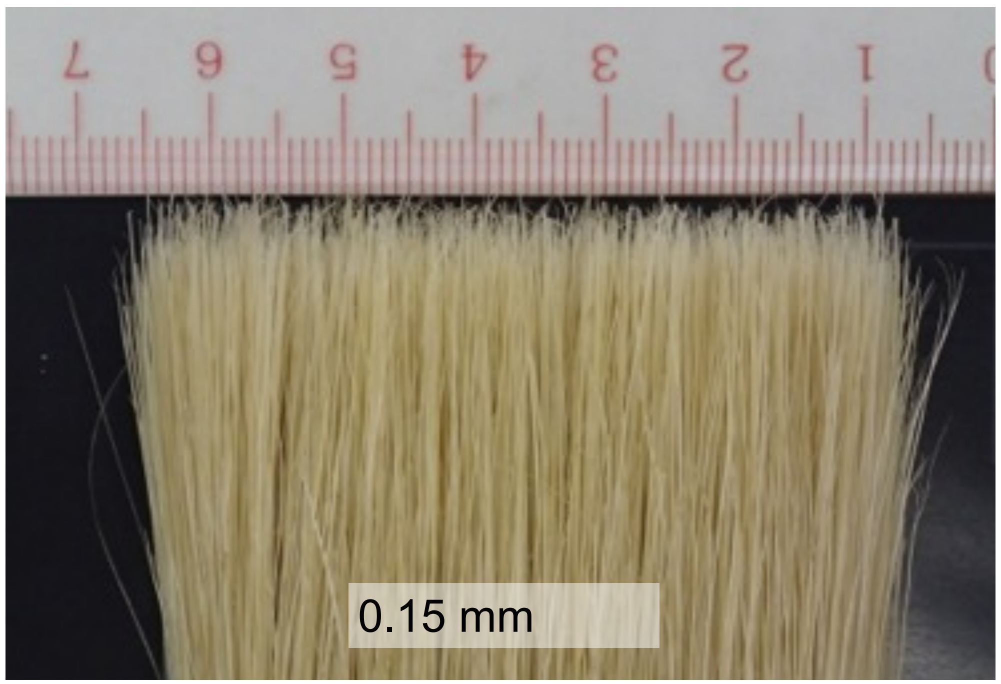


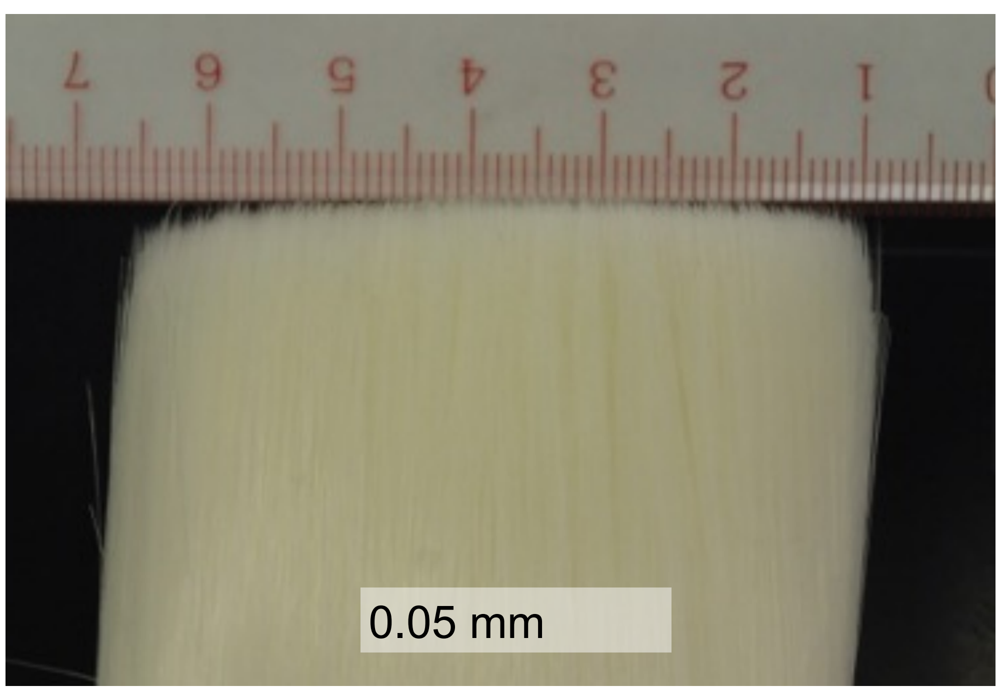
Fig. 3. Conceptual diagram of our band-sifting space. Input image courtsey of Ewan [2009] Given an image, we split it into high- and low-frequency subbands. These are then split into high- and low-amplitude parts. These are further split into positive and negative parts. For visualization purposes we show only two frequency splits, but in practice we create $\log_2(\min(\mathit{width}, \mathit{height}))$-frequency subbands and work on each one of them. Further, in order to make the size of the space more tractable, we "compress" the set of possible choices by looking at two categories of frequencies. We consider the high-to-mid frequencies as one category, which we refer to as "high-spatial frequencies", and we look at the mid-to-low frequencies as another category, which we call "low-spatial frequencies". Further, as shown in the diagram, we do not split the low-spatial frequencies category based on the amplitude of the coefficients since, numerically, sifting based on this criterion does not give much differentiation. However, the sign of the coefficients is still a useful sifting criterion along the low-spatial frequencies paths, for instance, it differentiates between broad-gloss and broad-shadow effects. With colored text and arrows we show how various operators can be mapped into paths in our space. With purple borders we show the path of sifted coefficients that was used to generate the orange result in Figure 1. In the upper-left border we show an alternative, more compact diagram, of the same path. On the far right, next to each path, we show acronyms which we use to describe paths in our space. For example, we use HHP for paths that manipulate (H)igh-spatial frequency, (H)igh-amplitude, (P)ositive coefficients.

Fineness

Perception of sub-resolution fineness

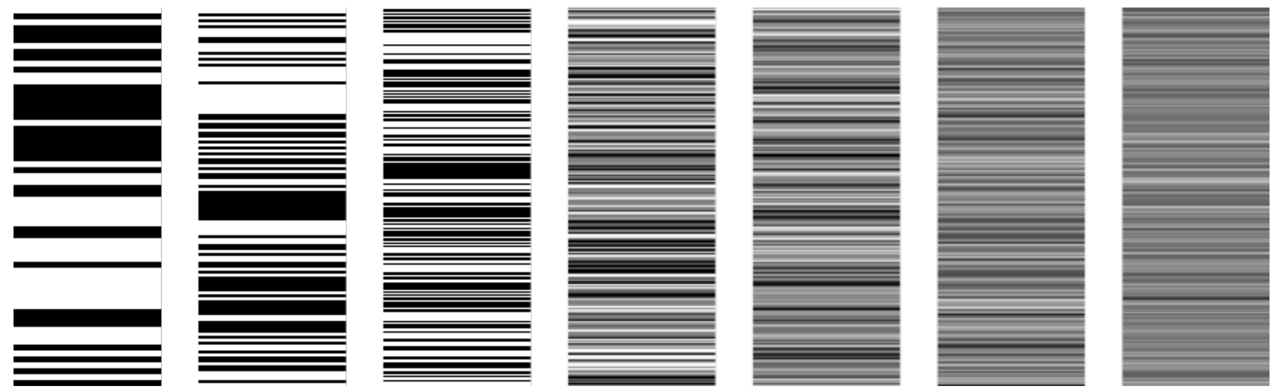
Shinya, Sawayama & Nishida (in preparation)



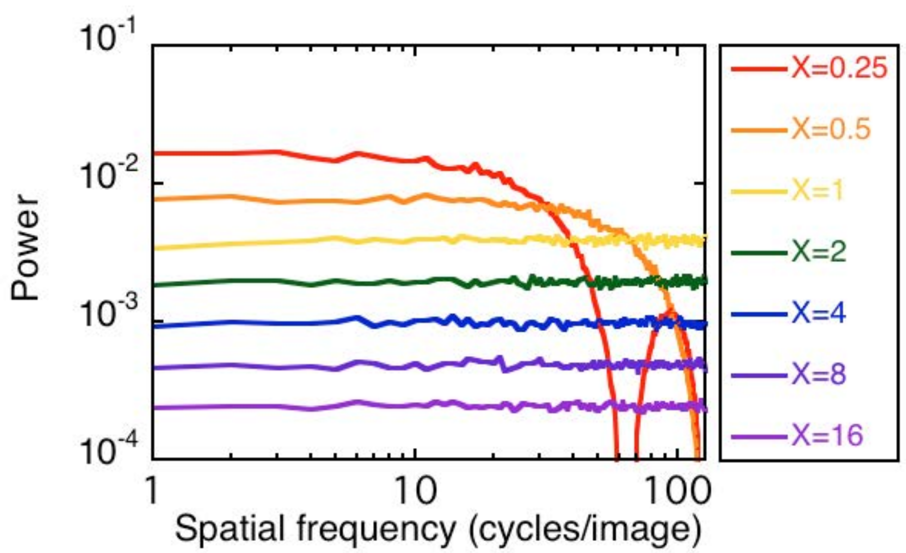


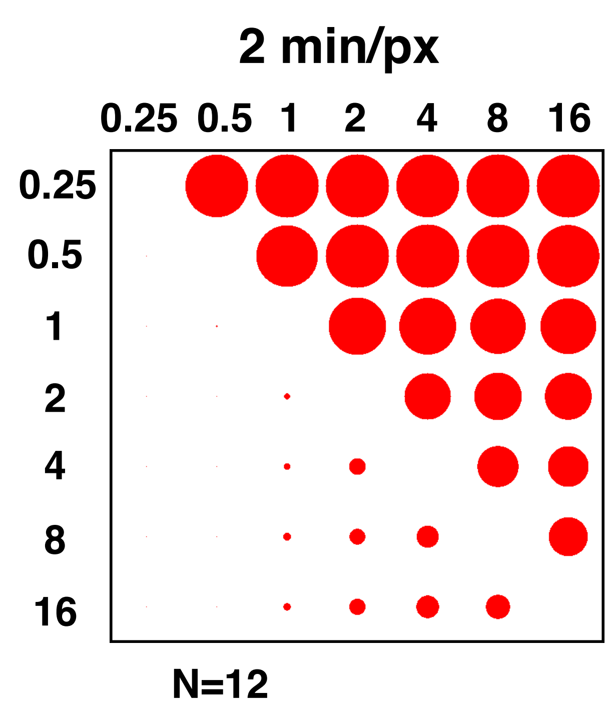
Perception of sub-resolution fineness

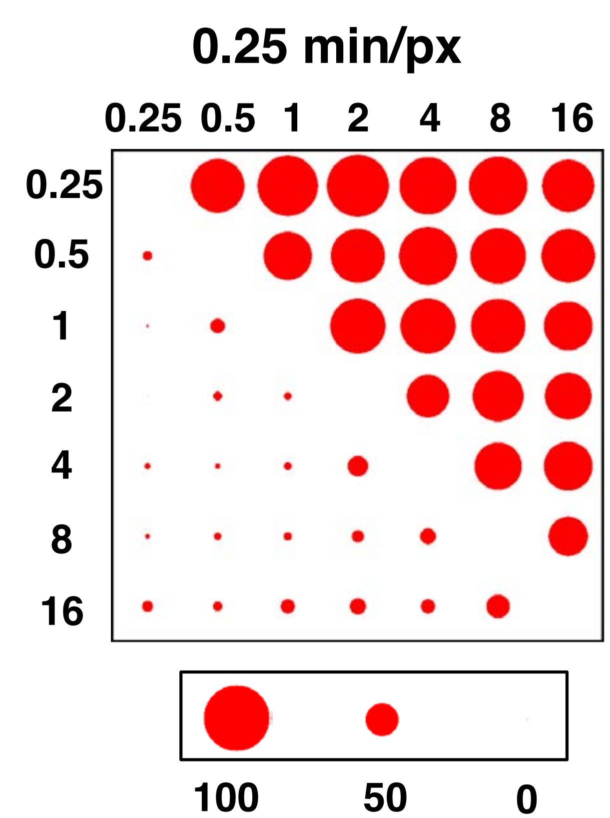
Shinya, Sawayama & Nishida (in preparation)



X (lines/pixel)	0.25	0.5	1	2	4	8	16
Lines/min at 43 cm distance	0.125	0.25	0.5	1	2	4	8
Lines/min at 344 cm distance	1	2	4	8	16	32	64

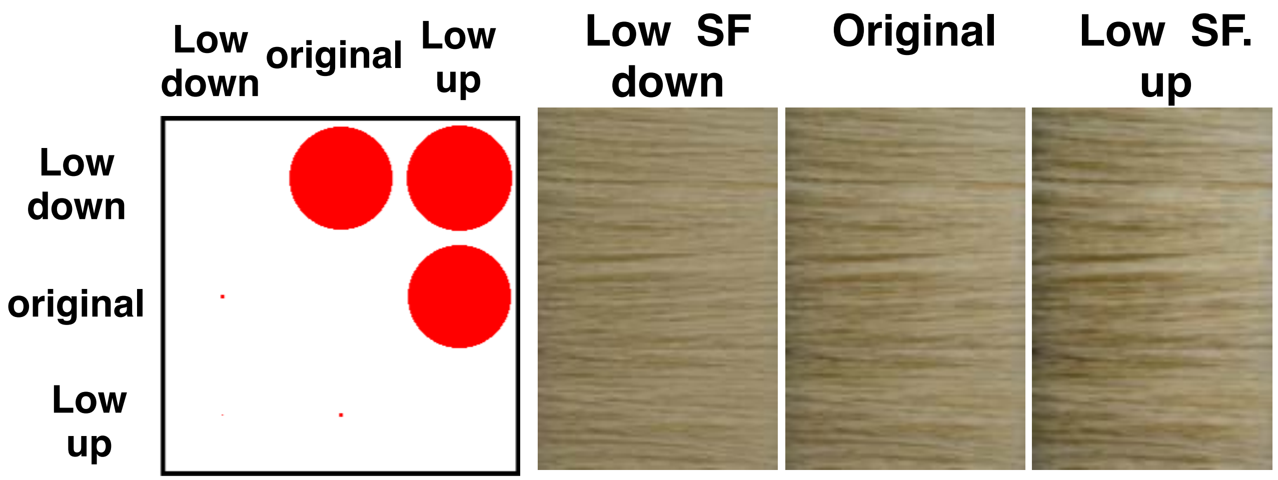




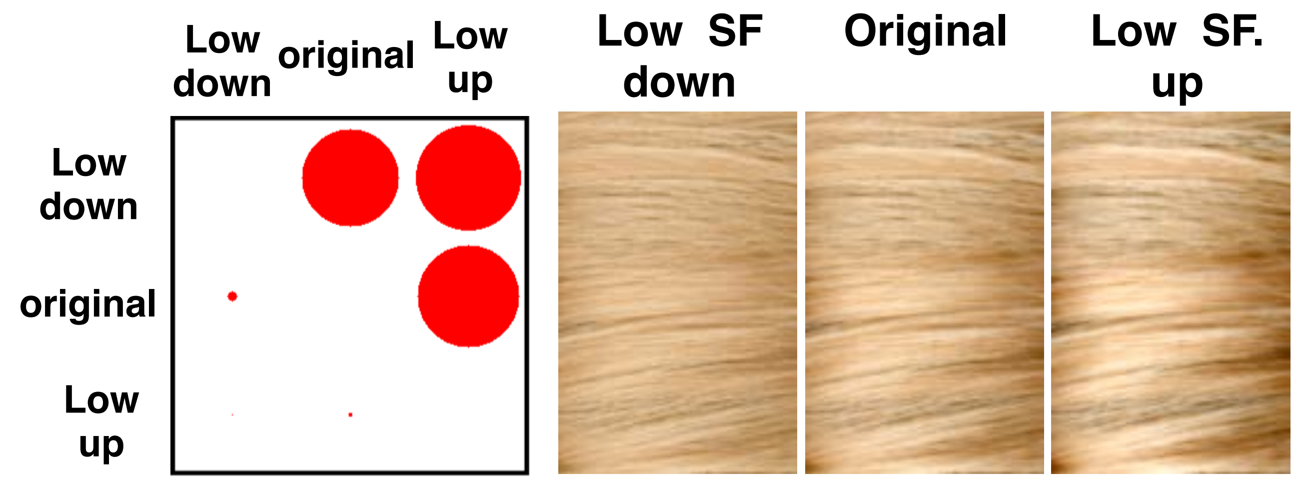


Perception of sub-resolution fineness

Brush

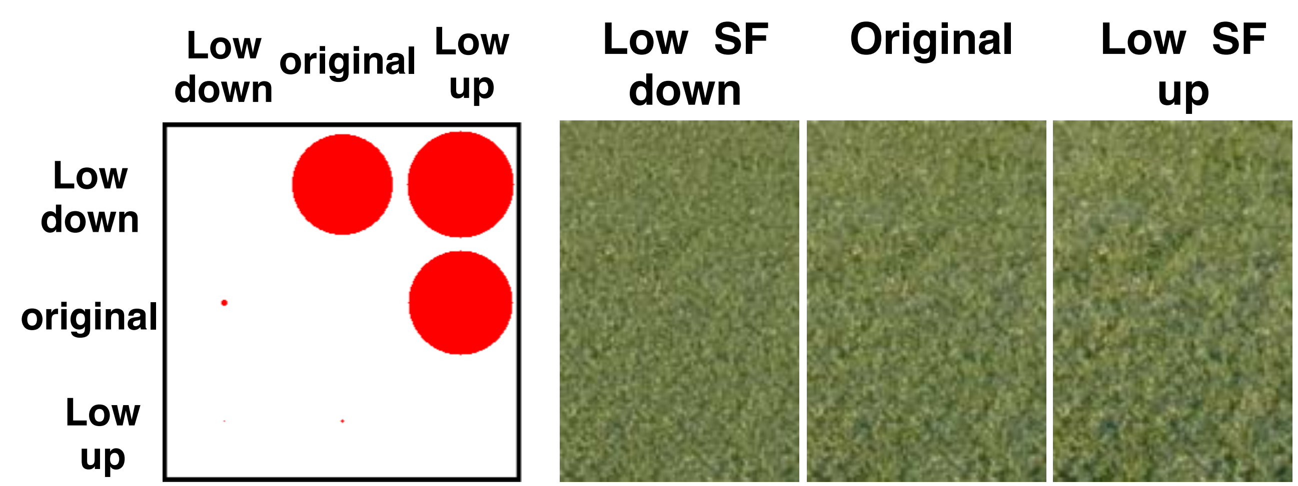


Hair 2



Hair 1

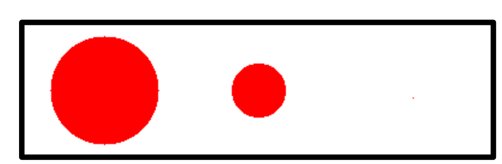
Grass



N=11

Shinya, Sawayama & Nishida (in preparation)

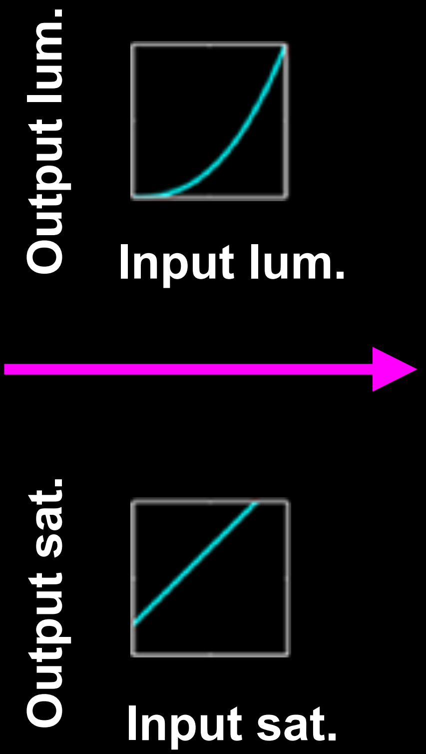
Percentage of trials in which “row” textures were judged to be finer than “column” textures



Wetness

Wet filter

Original image



Transformed image

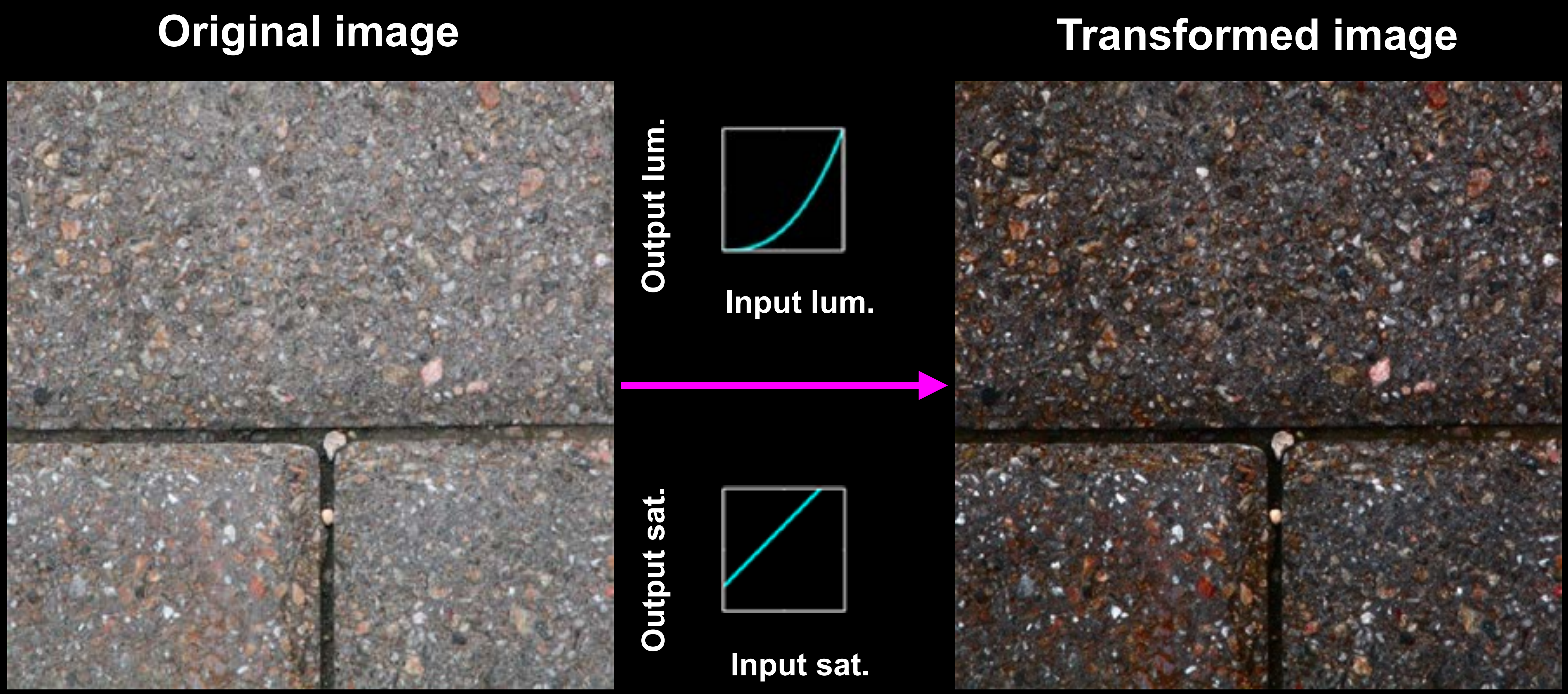
Operations

①Change the luminance histogram positively skewed

②Enhance the color saturation

Sawayama & Nishida (2015 VSS)

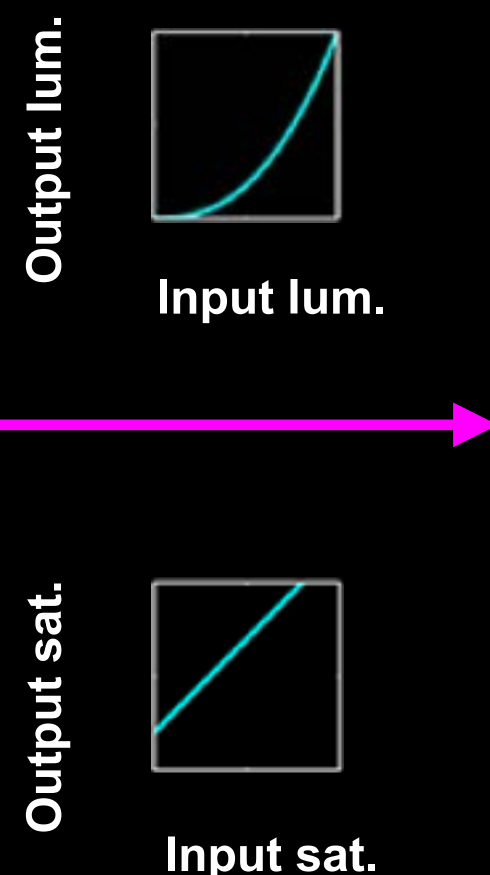
Wet filter



Original image
Transformed image
Output lum.
Input lum.
Output sat.
Input sat.

Wet filter

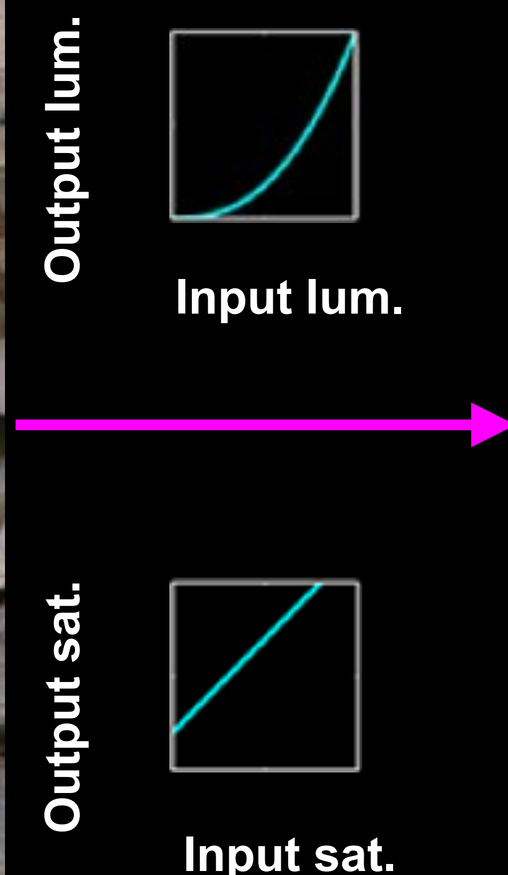
Original image



Transformed image

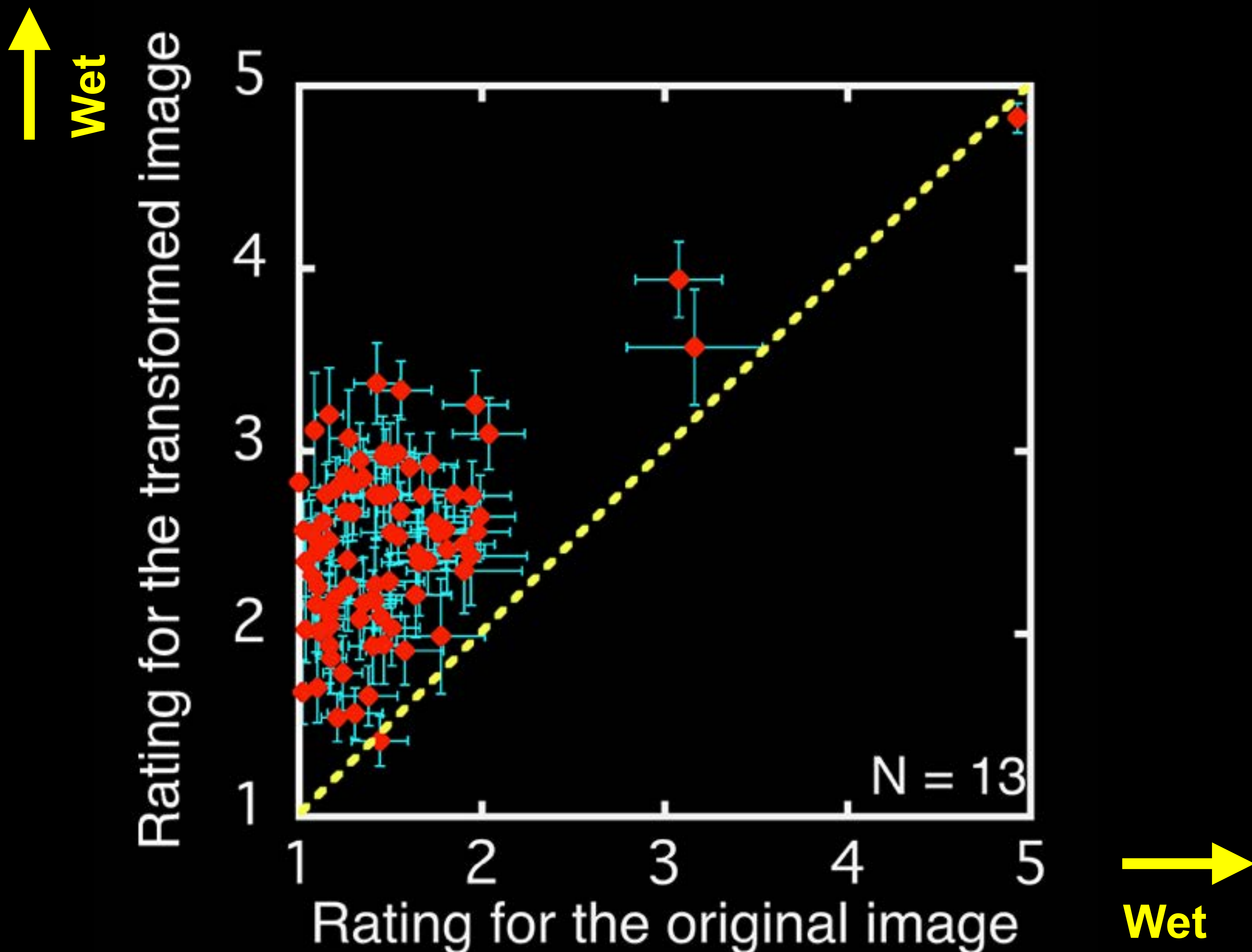
Wet filter

Original image



Transformed image

Results of wetness rating



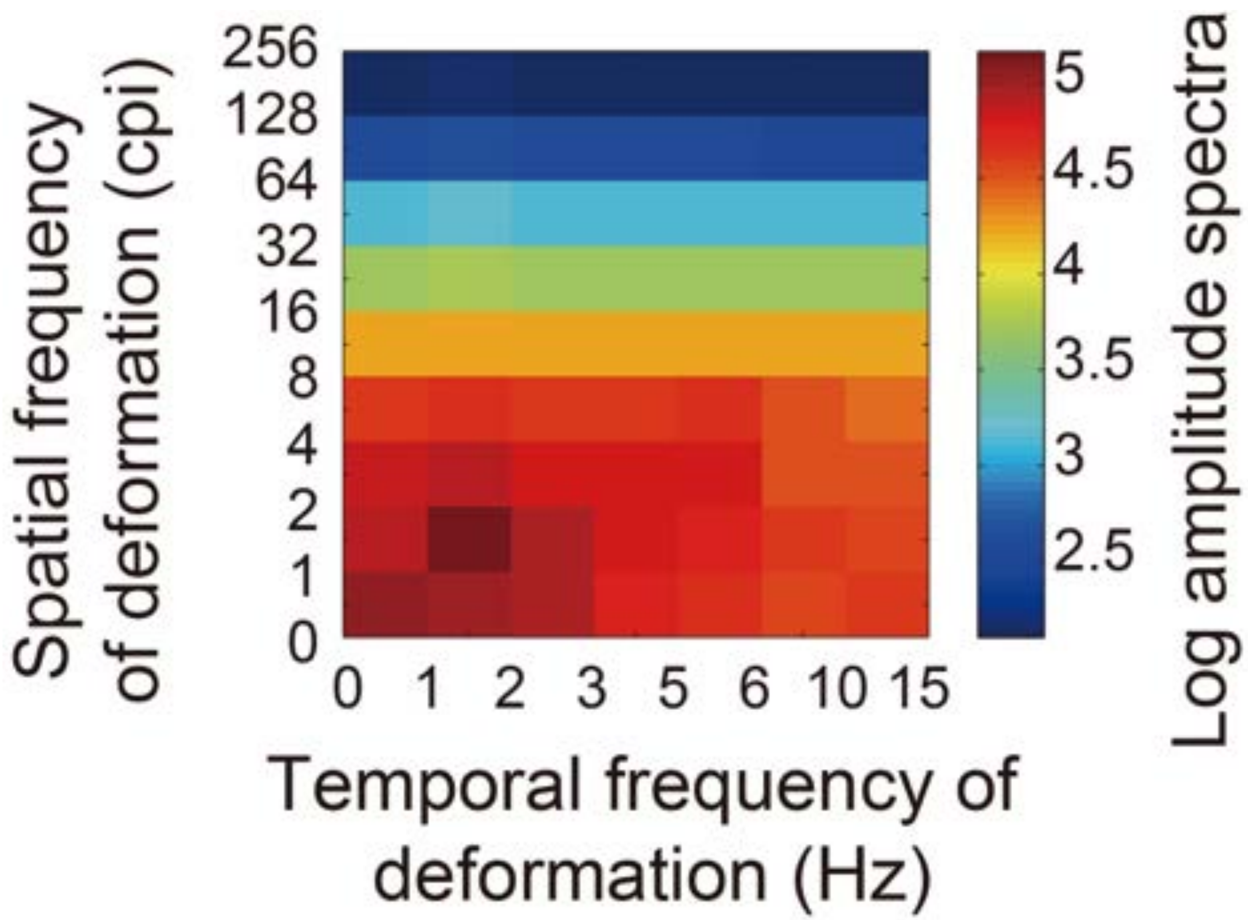
Transparent Liquid

Seeing liquid from image deformation

Kawabe, Maruya, & Nishida (2015, PNAS)

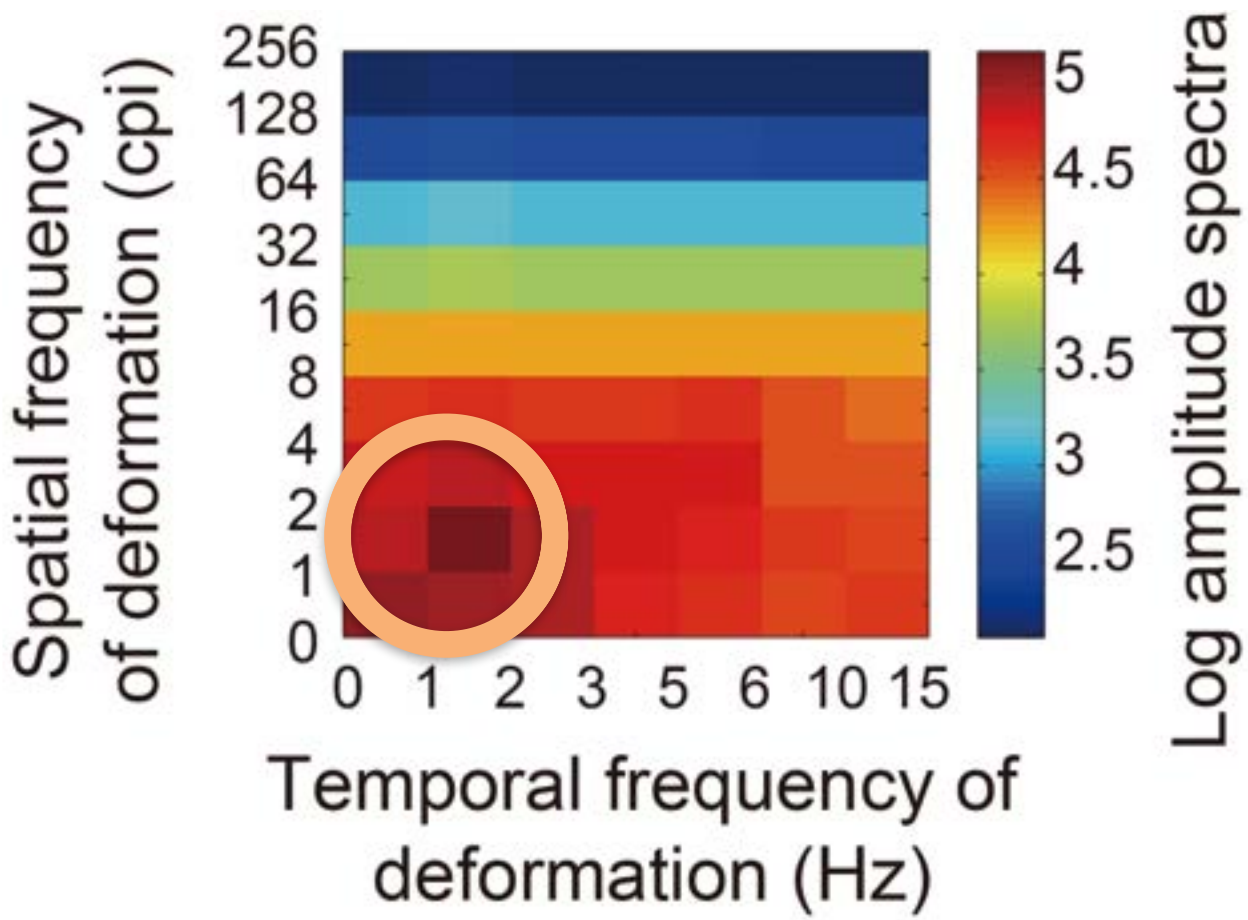
Spatiotemporal deformation frequency

Low-pass

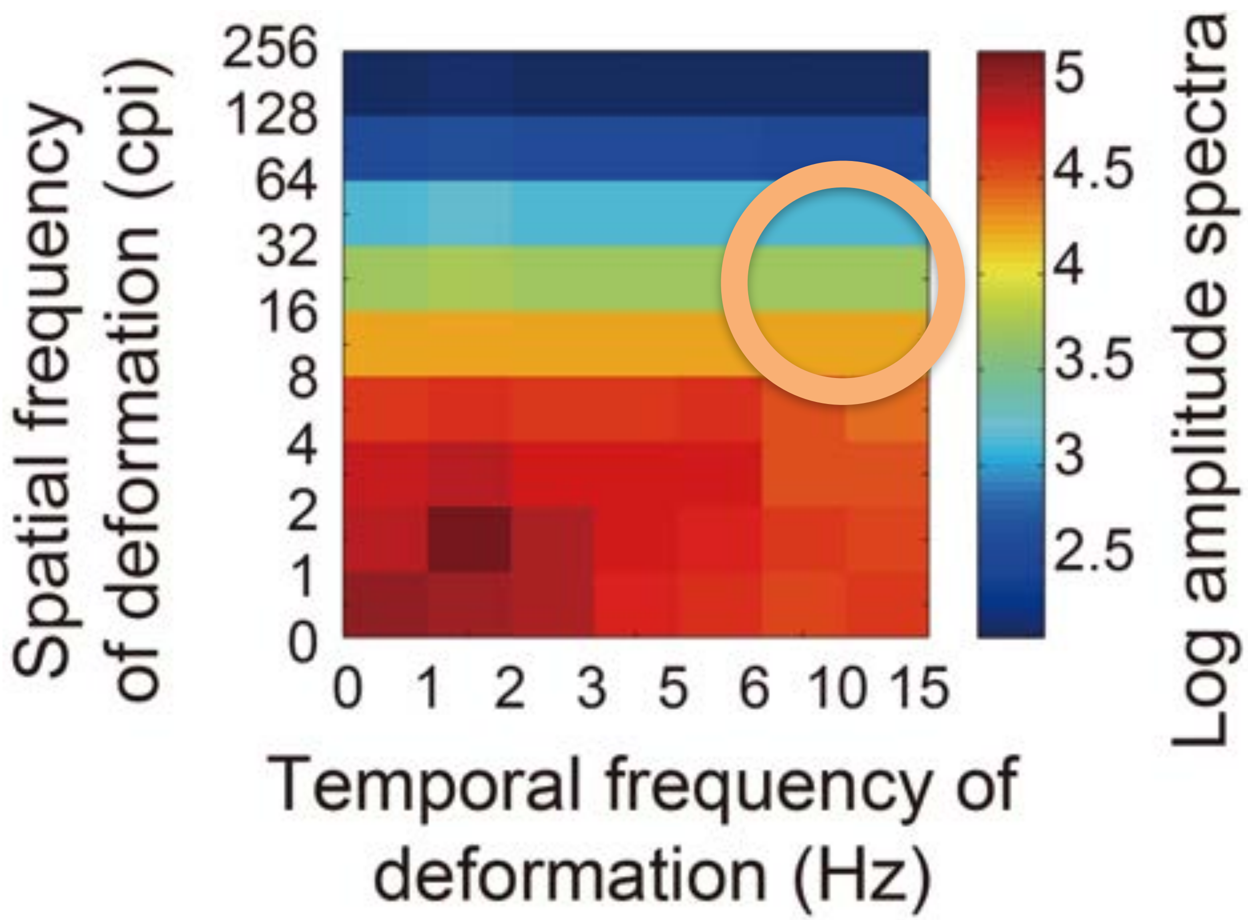


Broadband

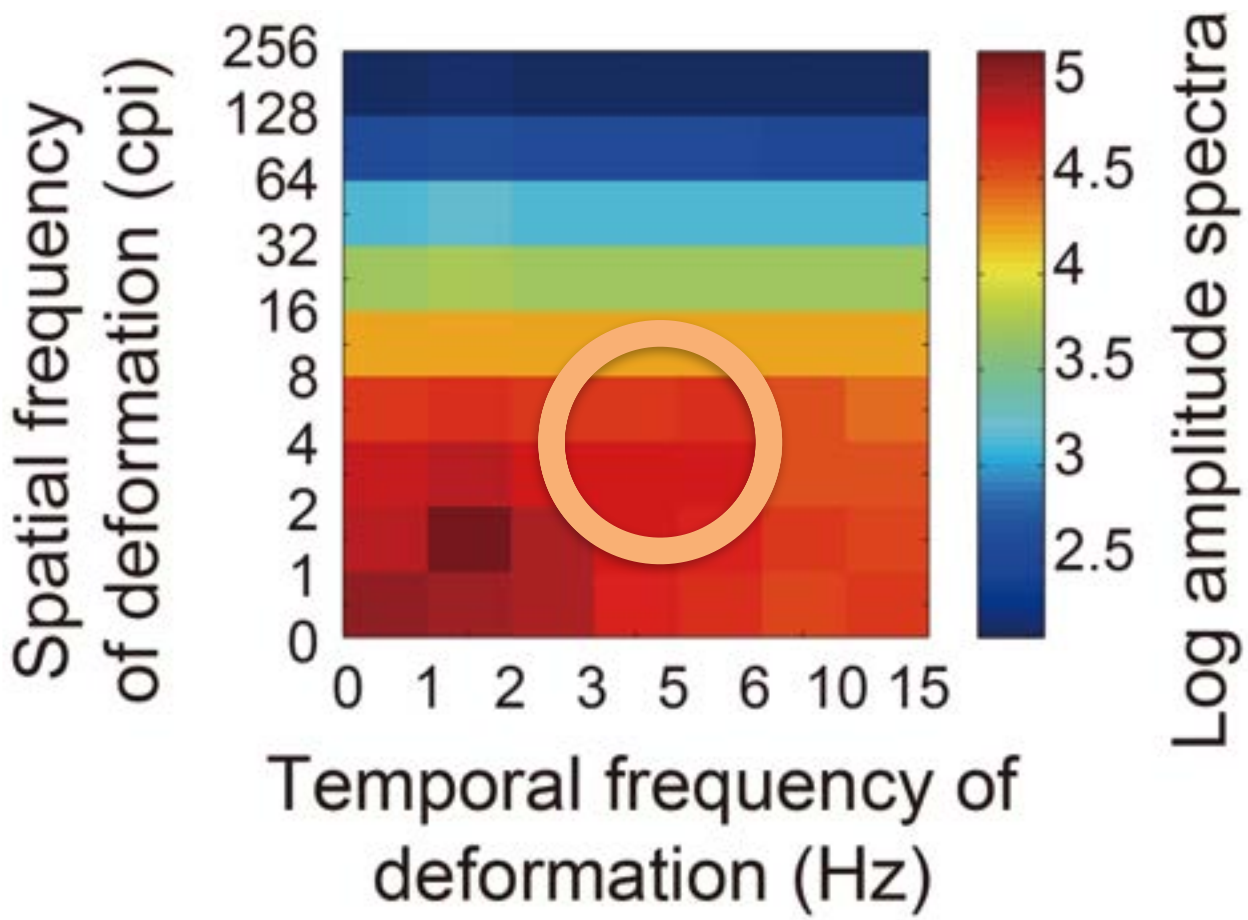
When is image deformation effective?



When is image deformation effective?



When is image deformation effective?

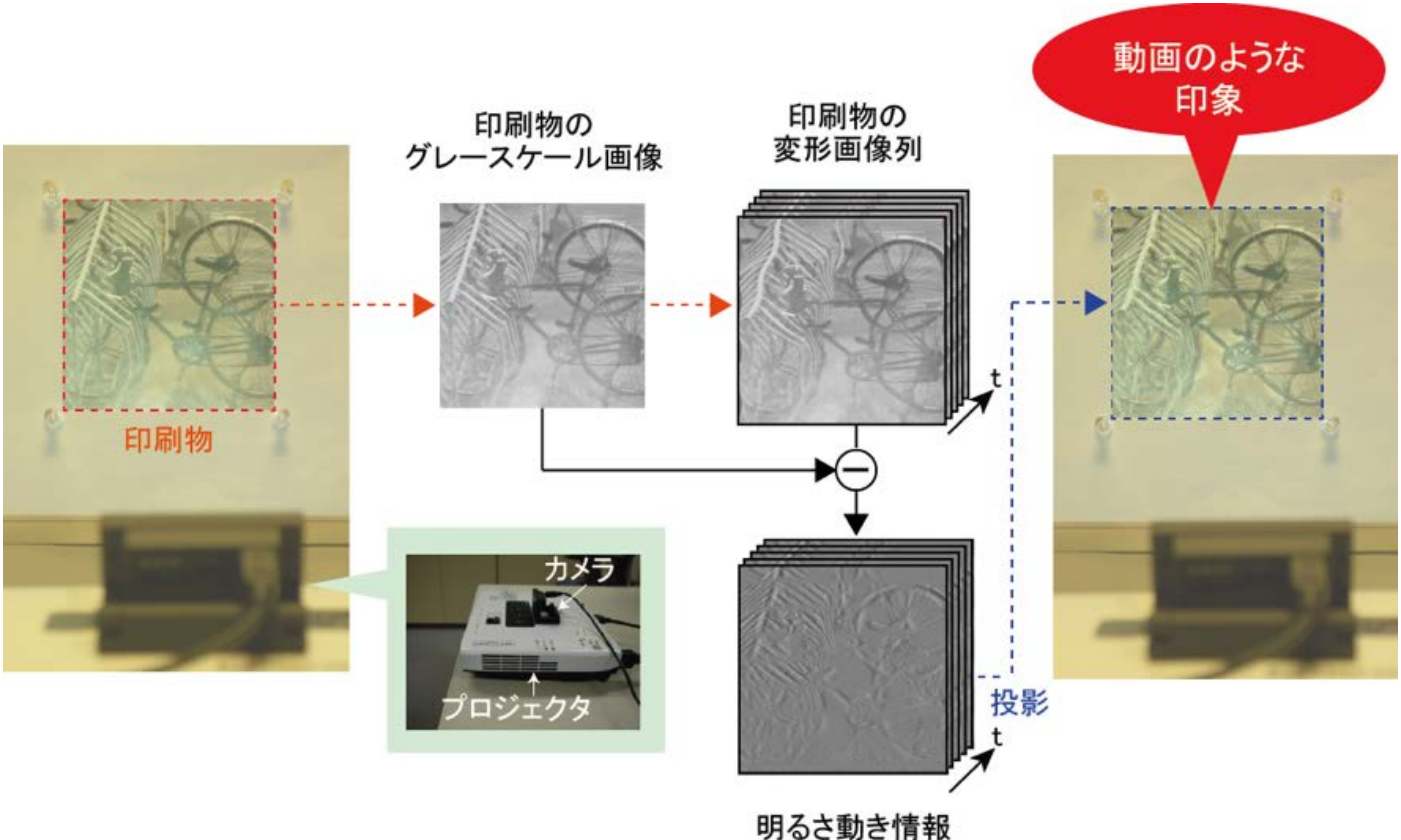


炎

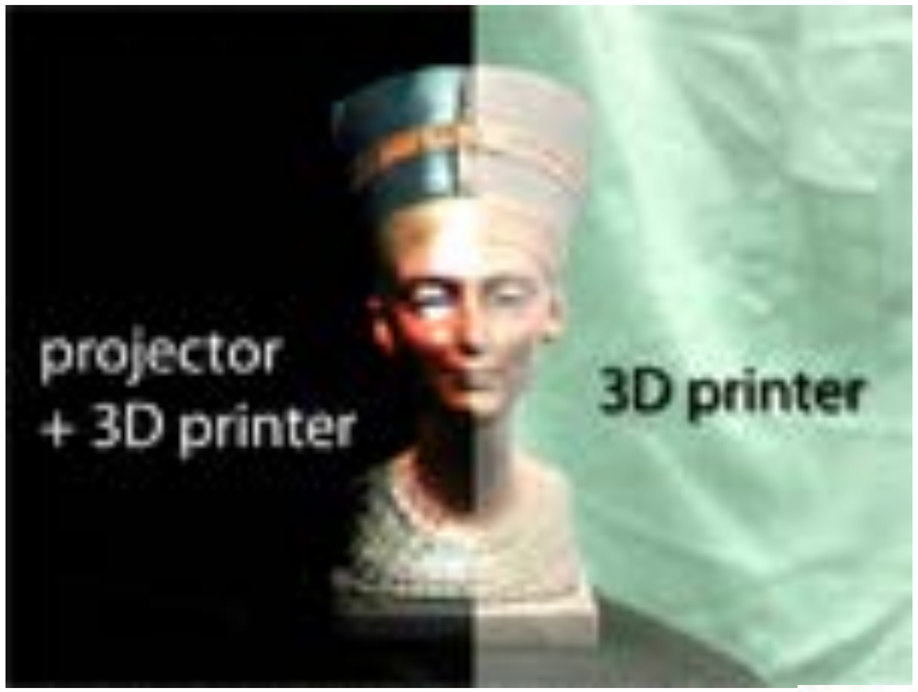
水

変幻灯システム：カメラ＋プロジェクタ

任意の静止画を様々に変形し、多様な効果（印象）を付加する



プロジェクションによる見かけの編集



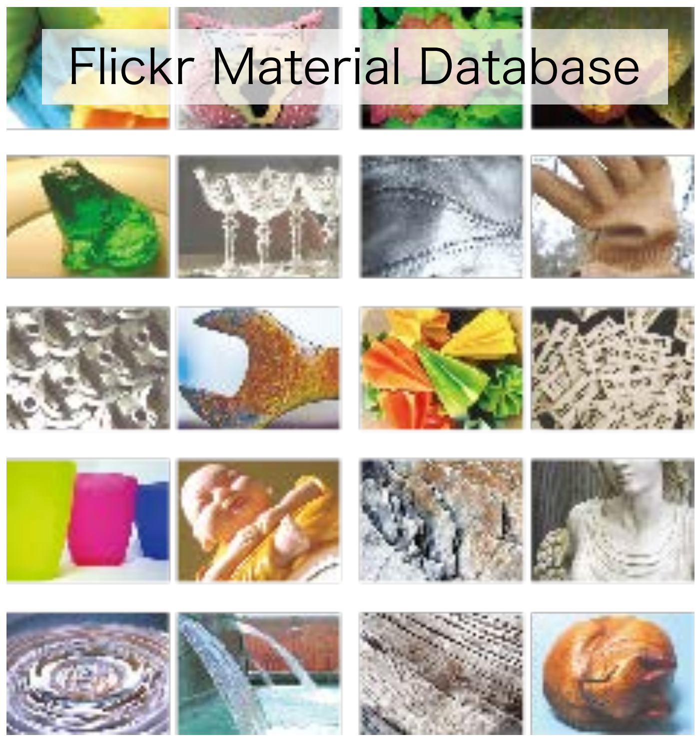
ダイナミックレンジ拡大(Bimber & Iwai, 2008; Shimazu, Iwai & Sato, 2011)

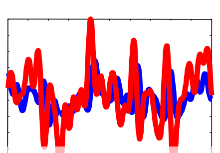
アピアランス制御 (Amano & Kato, 2008;天野・加藤, 2010)

実世界の多様な質感入力と質感出力

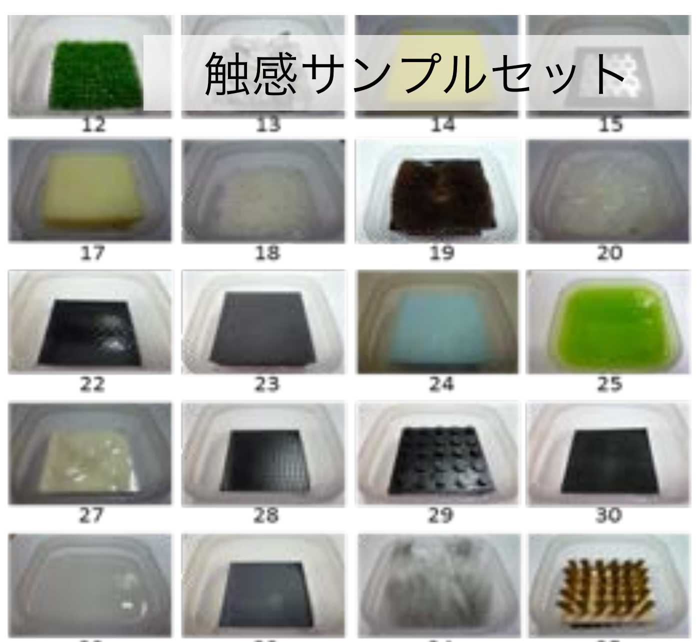
入力：質感刺激

質感サンプルセット





音響サンプル



出力：質感・意味・価値

光沢感　べたべた

粘性感

透明感

液体感

固体感

素材感

高級感

シズル感

情動

気配

音質

生き物らしさ

触感

温感

カワイイ

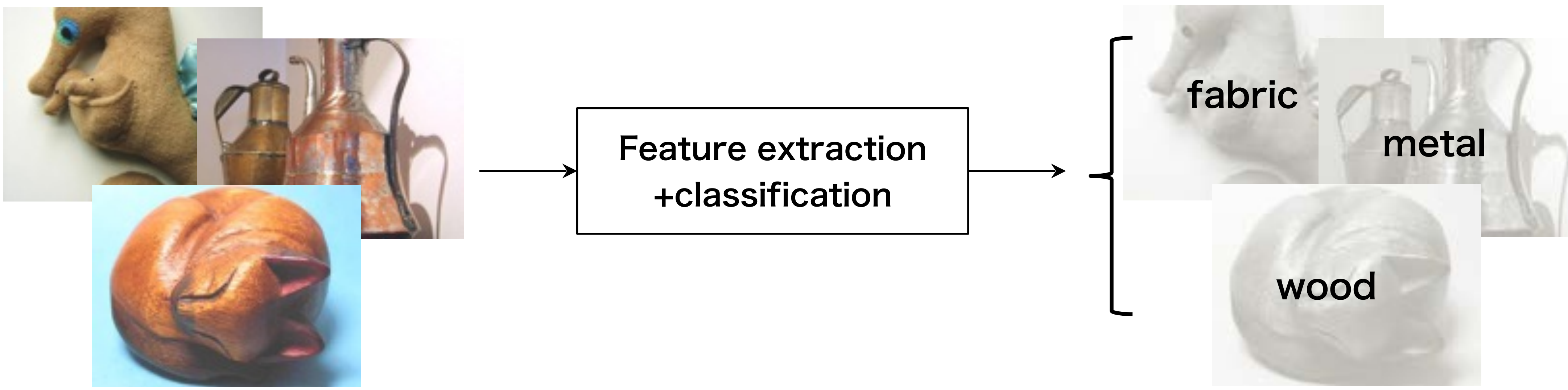
新しい情報科学・脳科学の流れを取り込んで

実世界の多様な質感を効率的に解明

多元質感知岡谷班キックオフスライドより

材質の画像認識

[Liu+2010][Hu+2011][Timofte+2012][Sharan+2013]…



多元質感知岡谷班キックオフスライドより

材質の画像認識

例：Fablic

多元質感知岡谷班キックオフスライドより

材質の画像認識

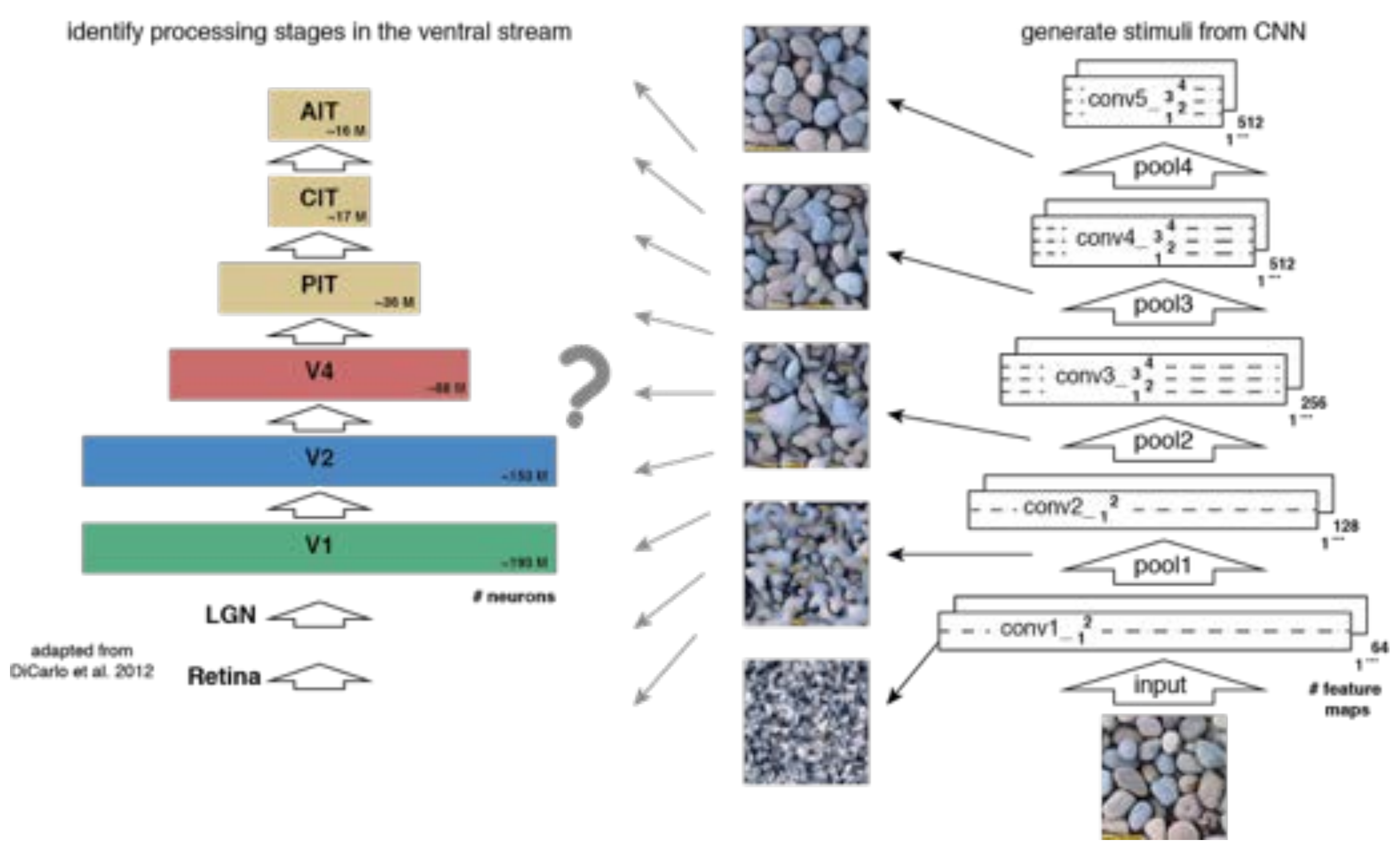
Liu, Ozay, Okatani, preparation for submission

Publication	Low level features	High level feat	Classifier	Recognition accuracy
Liu+2010	SIFT, color, curvature, edge-ribbon etc.	BoF	LDA	44.60
Hu+2011	Kernel descriptors(five)	Efficient match kernel	SVM	54.00
Timofte +2012	BIFs(Basic Image Features)	SPM-like + Comparative reasoning	CRC (Collaborative Representation Classifier)	55.78
Sharan +2013	SIFT, color, curvature, edge-robbon etc.	BoF	SVM	60.6
↑	Human (Mechanical Turk)	?	?	84.9
Cimpoi +2015	Deep CNN(VGG-19layers)	←	←	82.5±1.5

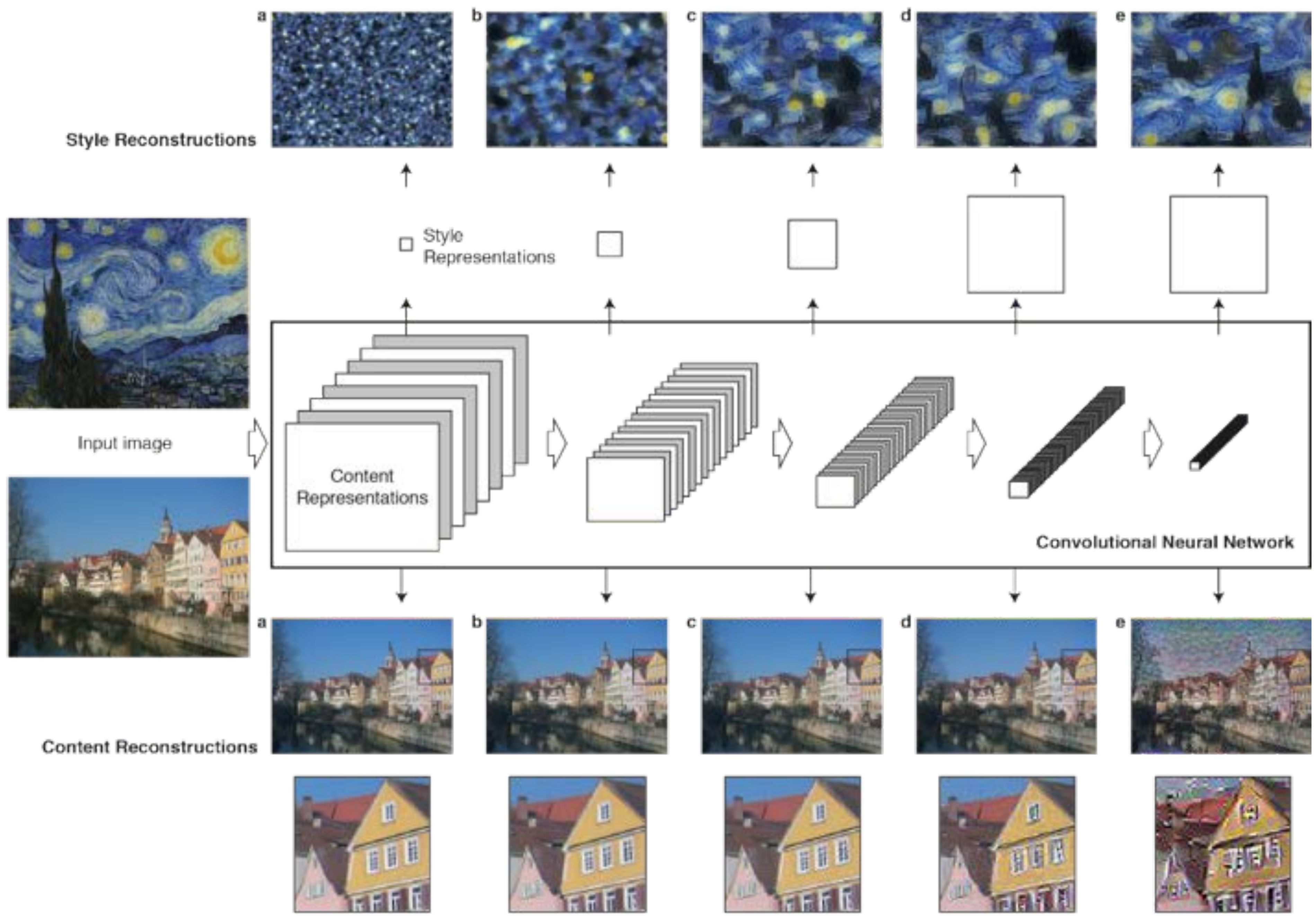
Ours (preparation for submission)	Ensemble of multiple Deep CNNs	←	Hierarchical sparse decision fusion	84.3±1.9

CNN-based texture synthesis

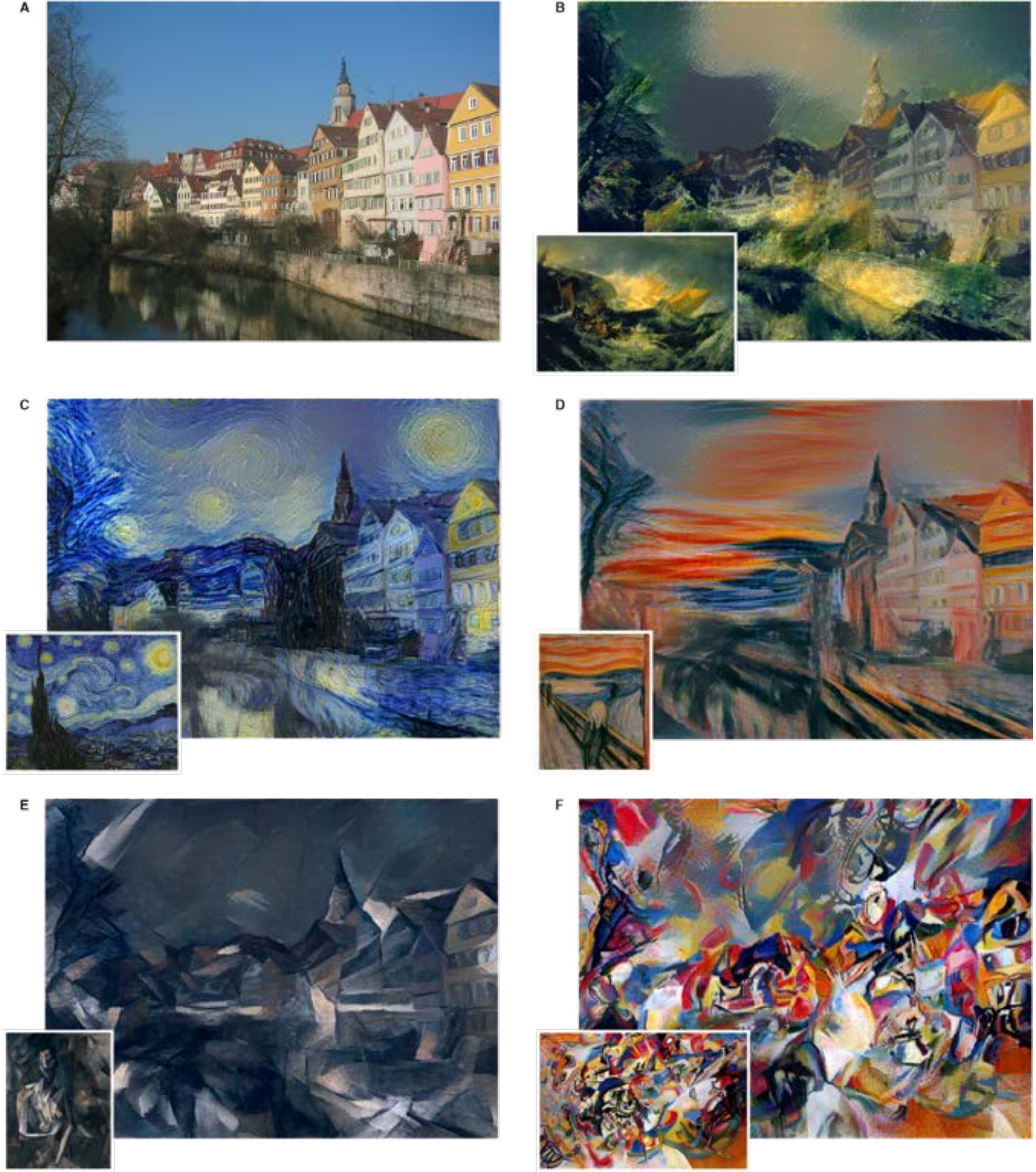
Gatys, L. A., Ecker, A. S., & Bethge, M. (2015). Texture synthesis and the controlled generation of natural stimuli using convolutional neural networks. arXiv.org.



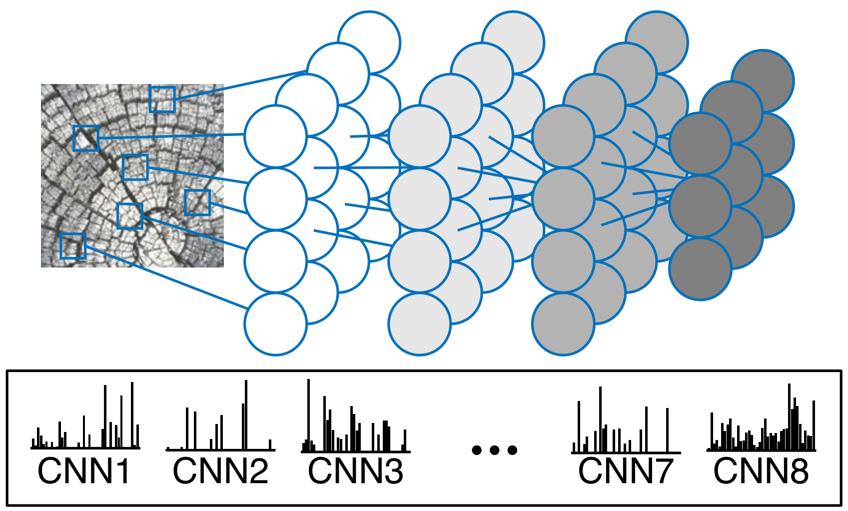
CNN-based texture editing

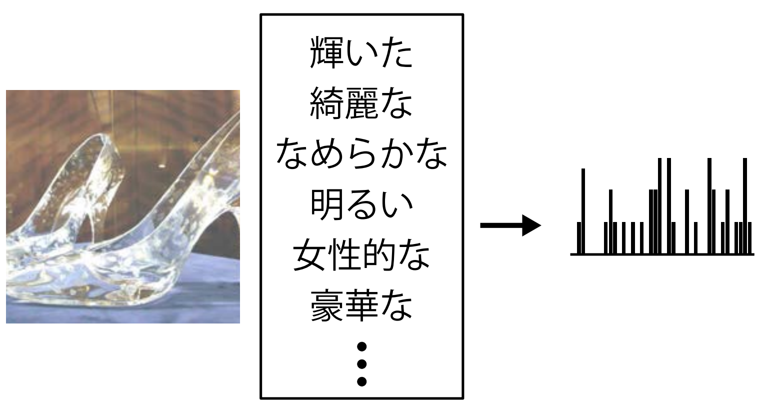


Gatys, L. A., Ecker, A. S., & Bethge, M. (2015). A Neural Algorithm of Artistic Style, arXiv:1508.06576 [cs.CV],1–16.

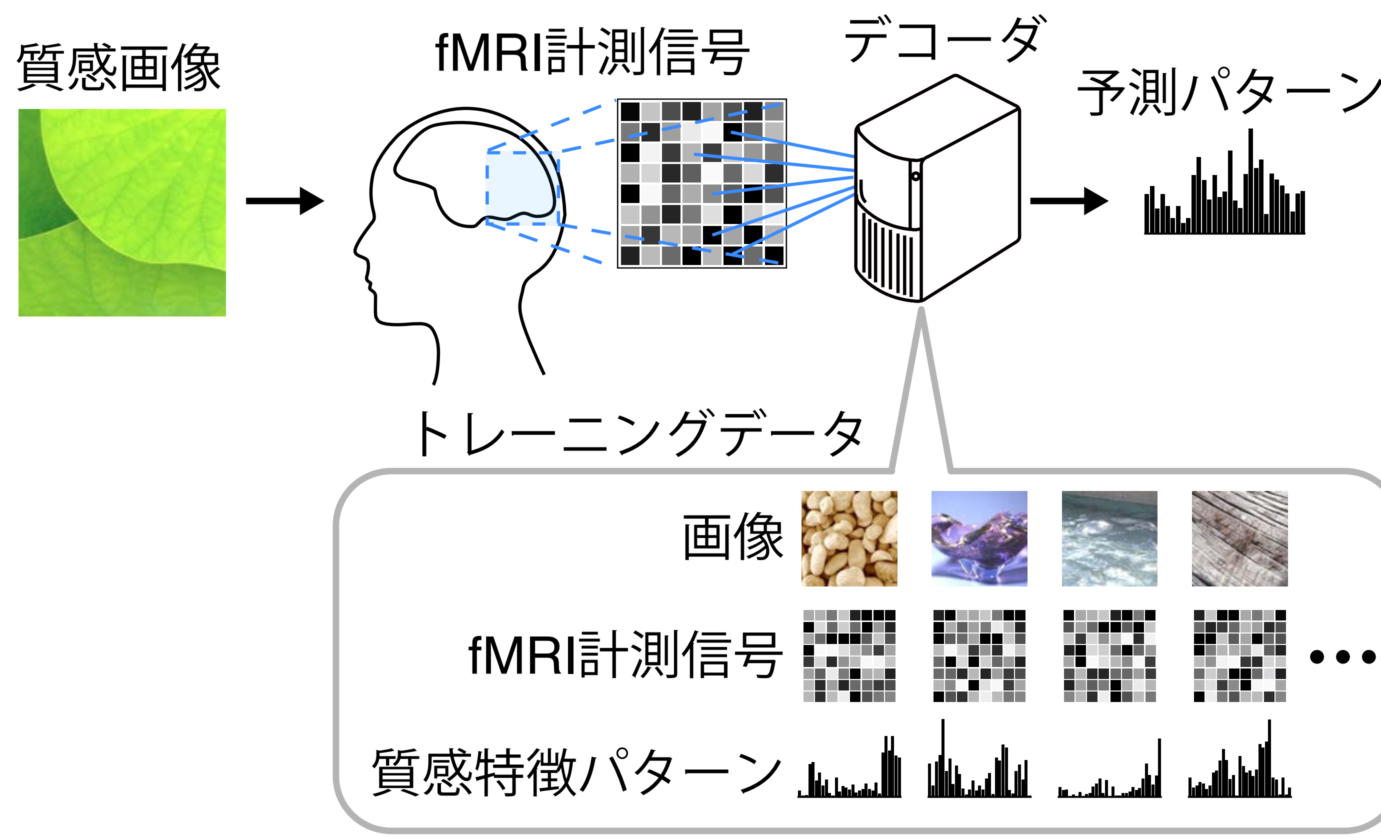

A
B
C
D
E
F

画像・テキストのデータマイニング技術と
デコーディング技術の連携による質感情報表現の探索





**ニューラルネットワークによる
画像・テキストデータ解析に
もとづく質感特徴抽出**



**fMRI計測信号からの
質感特徴予測を介した
脳内質感表現の探索**

最後に

- 脳の質感情報処理は浅すぎず・深すぎず
- 多様で複雑な質感の認識・操作
  - 理論検証型アプローチ
    - 着実に進歩しているが、問題は無限に広がっている
    - 研究方法論も模索中
  - データ駆動型アプローチ
    - Deep learning (CNN)の衝撃
    - 理解するには不満
    - 環境に情報がある