

マルチピンホールカメラにおけるプライバシー保護指標の検討

田中滯士¹ 佐藤智² 石井育規²

概要: 顔認証や行動認識技術に対してプライバシーに関する問題が懸念されている。我々が提案するマルチピンホールカメラは、これらの技術における高い認識精度とプライバシー保護を両立できるが、プライバシー保護性能を評価する指標がなかった。本報告では、マルチピンホールカメラのプライバシー保護性能を評価するために、画像からの人物特定容易性の指標である秘匿性評価指標を導入する。さらに、マルチピンホールカメラで撮影される画像の秘匿性を客観的に求めるため、Mean Squared Error (MSE)、Structure Similarity (SSIM)、Most Apparent Distortion (MAD)の3つの画質評価指標を、①主観的な秘匿性評価との類似性、②計算時間、の2点で比較・評価し、MSEが最も実用的な評価指標であることを確認した。

キーワード: マルチピンホールカメラ, プライバシー保護, 秘匿性評価, 画質評価

Examination of privacy protection index for multi-pinhole camera

REIJI TANAKA¹ SATOSHI SATO² YASUNORI ISHII²

Abstract: There are concerns about privacy protection for face recognition and human activity recognition technologies. The Multi-pinhole camera we propose can achieve both high recognition accuracy and privacy protection in these technologies. However, we have no index to assess the degree of privacy protection. In this report, in order to assess it, we introduce secrecy assessment index which represents difficulty of identifying a person from the image. Furthermore, for objectively determining the secrecy assessment index of an image taken by multi-pinhole camera, we compared and evaluated three image quality assessment indexes: Mean Squared Error (MSE), Structure Similarity (SSIM) and Most Apparent Distortion (MAD), in regards to (1) the similarity with subjective assessment and (2) calculation time. As a result, we confirmed that MSE is the most practical assessment index.

Keywords: Multi-pinhole camera, Privacy protection, Secretary assessment, Image quality assessment

1. はじめに

顔認証や行動認識は様々なシーンで利用される。例えば、家庭内で個人や行動を認識することにより、ユーザに適したサービスを提供できる。しかしながら、撮影した画像は個人を特定し得る情報が含まれている。また、ユーザは家庭内の散乱している状態を他人に見られたくない。このように、撮影場所によってはプライバシーに関する問題がある。そのため、プライバシーを保護した画像のセンシングが強く求められている。

プライバシーの保護と画像認識精度を両立するためのデバイスとして我々はマルチピンホールカメラ (Multi-Pinhole Camera: MPC) を提案している [1]。このカメラは、イメージセンサの前に複数個の穴 (ピンホール) を開けたマスクが配置されている。そのため、撮影画像は複数のピンホールから入射した光が重ね合わされ、視点の異なる複数の画像を加算した像となるため、画像に何が写っているのかを人の目で正確に判別することは難しい。また、既存の画像復元技術ではピンホールの配置が未知では鮮明な画

像を復元することは難しい [1]。このように、MPCはプライバシーを保護した画像を撮影できる。

MPCによりプライバシーは保護できるが、重なり方によっては認識精度の低下を招く。つまり、マルチピンホールの配置と認識精度は密接に関係しており、トレードオフの関係にある。したがって、要求される認識精度とプライバシーの保護を両立可能なマルチピンホールマスクを設計する必要がある。

マルチピンホールマスクを最適に設計するためには、プライバシーを保護できるか否かのプライバシー保護性能を評価する客観的な指標が必要である。この指標は、人が見てプライバシー保護性能が高い場合は高くなり、人が見てプライバシー保護性能が低い場合には評価値が低くなることが期待される。しかしながら、MPCで撮影した多重像のプライバシー保護性能の定量化については、これまで検討されておらず、プライバシー保護性能を表す評価指標はなかった。

そこで本稿では、MPCで撮影した多重像を対象として、人が見たプライバシー保護性能と同様の傾向を有する客観

¹ 大阪大学大学院生命機能研究科
Graduate School of Frontier Biosciences, Osaka University
² パナソニック株式会社
Panasonic Corporation



Figure 1. (Left) Normal image, (Right) Multi-pinhole image

指標について検討する。まず、多重像やノイズ付加などプライバシー保護性能を変えた画像をユーザに呈示し、主観的なプライバシー保護性能の順位付けを行う。次に、画質評価基準である Mean Square Error (MSE), Structure Similarity (SSIM) [2], Most Apparent Distortion (MAD) [3]を用いて画質に関する評価値を計算し、主観的なプライバシー保護性能と比較することで、最適な客観的評価指標を求める。

2. マルチピンホールカメラと秘匿性

2.1 マルチピンホールカメラ (MPC)

Fig. 1に、通常カメラとMPCで撮影した画像を記載する。通常カメラは、広義のピンホールカメラである。ピンホールカメラは、ピンホールと呼ばれる小さな穴を通過した光線のみをイメージセンサで受光し、合焦した画像を撮影する。ピンホールカメラでは、ピンホールを通過しない光線は捕捉されないため、シャープな画像を撮影できる。本来、カメラの目的は、明るくシャープな画像を撮影することである。しかしながら、このような画像では被写体の状態を容易に判別できるため、家庭内などプライバシーに配慮すべき環境での利用には適さない。

MPCは、イメージセンサの前に複数個の穴(ピンホール)を開けたマスク(マルチピンホールマスク)が配置されているため、撮影画像は複数のピンホールから入射した光が重ね合わされたものとなる。このカメラを用いて撮影した画像は、視点が異なる複数の画像を加算した像となるため、画像に何が写っているのかを人の目で正確に判別することは難しい。このように、MPCを用いることでプライバシーを保護した画像を撮影できる。

2.2 マルチピンホールカメラの秘匿性

マルチピンホールマスクはピンホールの数、サイズと位置を設計して作成される。認識精度とプライバシー保護性能はマルチピンホールの設計に依存する。そのため、最適なマルチピンホールマスクを設計するためには、認識精度とプライバシー保護性能を評価する必要がある。また、このプライバシー保護性能は、撮影画像のみならず、撮影画像が流出したときのプライバシー漏洩リスクの考慮が必要である。本稿で述べるプライバシー漏洩リスクとは、画像流出時にプライバシー保護していない画像に画像復元され

るリスクを意味する。すなわち、プライバシー保護性能の評価値は、撮影画像と復元画像の両方を評価できる必要がある。本稿では、運用中のMPCからマスクを解析することは困難であると考え、マルチピンホールマスクの形状は未知として、画像を復元し、評価する。

また、家庭内での利用を考えた場合、プライバシー保護としては「画像に写っているのは誰か」を人が見て特定できるかどうかが重要である。さらに、MPCを用いた顔認証や行動認識技術は様々な場面での利用を想定しているため、開発において莫大な量の画像について、プライバシー保護性能を確かめる必要がある。

そこで、我々は、MPCにおけるプライバシー保護性能を評価するために、以下の条件を満たす必要があると判断した。

1. 画像から人物が特定できるかできないかの指標であること
2. MPCでの撮影画像に加え、ピンホールの設計が未知で復元された画像に対しても利用できる指標であること
3. 1.および2.を満たす主観評価指標と一致する客観指標であること
4. 処理時間も考慮した指標であること

これらの条件を満たす客観的指標を秘匿性、条件1.および2.を満たす主観評価値を主観的秘匿性と呼ぶ。

3. プライバシー保護のための主観的評価値

MPCの秘匿性は人が見て画像上の人物が特定できるか否かの主観的な評価に依存する。そのため、客観的指標を求めるためには、この主観的な評価値を取得する必要がある。そこで、我々は主観評価値を取得するための実験方法を提案し、その評価値を求めることで秘匿性の評価基準を求める。

3.1 主観評価値の取得方法

人が画像から人物を特定できるかできないかを判断する基準は、エッジや色など非常に多くの要素に依存していると推測される。したがって、個人によって評価の基準が異なり、1枚の画像を用いて複数被験者の評価値を比較するのは容易ではない。そこで我々は、全被験者が「画像対での人物特性しやすさの違い」という同一の尺度で評価できる一対比較法を利用し、主観評価値を取得する。

主観評価実験において被験者に呈示した画面をFig. 2に示す。画面上段にはボケのない通常画像、下段には2種類の異なる評価画像を並べる。この画面を呈示し、「人物を特定しやすい画像は左右どちらか？」と被験者に質問し、次の5段階で回答を要求する。

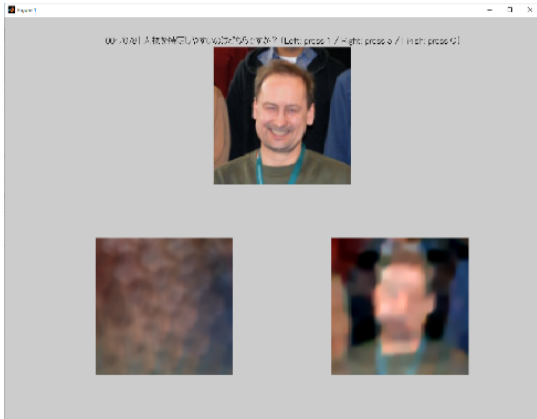


Figure 2. Task window presented to subjects

1. 左の方が明らかに人物を特定しやすい
2. 左の方が少し人物を特定しやすい
3. どちらとも言えない
4. 右の方が少し人物を特定しやすい
5. 右の方が明らかに人物を特定しやすい

このような5段階評価とすることで、画像間の秘匿性の差をより細かく評価できる。

被験者の回答結果より評価スコアを計算する。左に画像A, 右に画像Bを呈示した場合、各画像に対しTable 1に示した点数を加算し、被験者ごとに総和を求めることで評価スコアを得た。評価スコアが高い画像ほど人物特定が困難であり、評価スコアが低いほど人物特定が容易である。

3.2 評価した画像

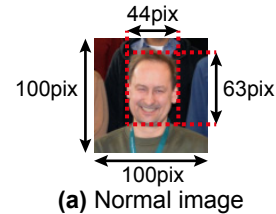
Fig. 3aに通常カメラで撮影される100×100 pixelsの画像[4]を示した。このうち顔領域は44×63 pixelsである。これを元画像として、以下の評価画像をシミュレーションで作成し、評価した。

- ・マスクパターンを変更してMPCで撮影した画像
- ・上記画像にノイズを付与した画像
- ・上記画像をBlind Deconvolution手法であるSelfDeblur法[5]で復元した画像

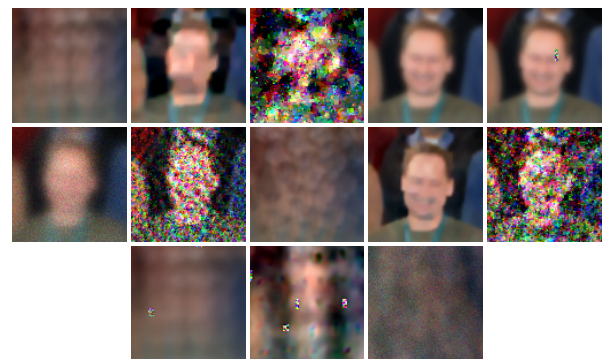
様々なマスクパターンで大量の画像を生成し、本稿ではその中からランダムに選択した13枚の画像(Fig. 3b)について評価を行う。主観評価実験においては、これら13枚の評価画像から、前項の通り2枚ずつ(およびFig. 3aの元画像)を被験者に呈示し、全ペア(78組)についてランダムな順番で評価を要求する。したがって、被験者に対して各評価画像は12回ずつ呈示されるため、各評価画像の評価スコアが取りうる値は最小12, 最大60である。この実験を11人の被験者で行い、各評価画像の平均評価スコアを主観的評価値とした。

Table 1. Conversion table from answer to score

Answer	1	2	3	4	5
Image A	+1	+2	+3	+4	+5
Image B	+5	+4	+3	+2	+1



(a) Normal image



(b) Test images

Figure 3. Assessed images

3.3 主観評価実験の結果

Fig. 4に、主観評価実験の結果を示す。この図において、人物を特定しやすいと評価された順に、左から右に画像を並べた。各評価画像の上部には生成パターンのインデックスを、下部には評価スコア(平均±標準偏差)を示した。評価スコアが低い画像は、顔の輪郭や目鼻の位置、髪型などが正確に推測でき、人物特定が容易である。評価スコアが高くなるほど輪郭等の識別は難しくなり、評価スコアが最も高い画像Lや画像Mは、人物の画像であるか否かさえ判別が難しい。つまり、評価スコアが高いほど、人物特定が困難である。したがって、本実験によって得られる各画像の評価スコアは、主観秘匿性の評価指標として有用であると言える。

また、Fig. 3に示した画像に加え、他の2人の人物画像についても同様に主観評価実験を行った結果をFig. 5に示す。それぞれの画像について、被験者3人の評価スコアの平均±標準偏差を示した。画像の生成パターンが同一の画像であれば、写っている人物が異なっても、主観評価は近い値となった。このことから、本実験は、写っている人物や背景に依存して結果が極端に変化する実験ではなく、様々な画像の秘匿性を主観的に評価するための汎用的な実験であることがわかった。

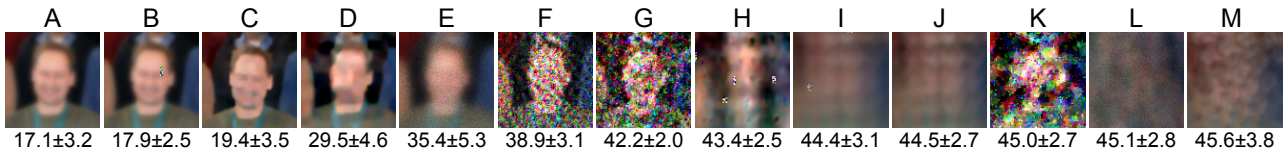


Figure 4. Result of subjective assessment

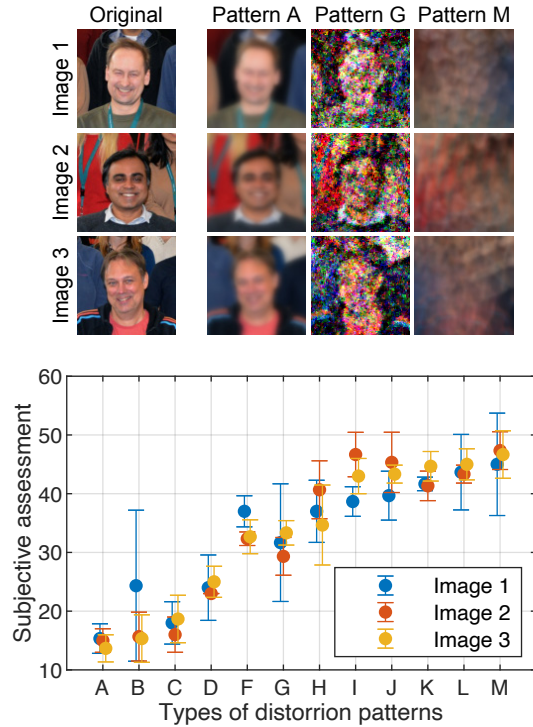


Figure 5. Differences in subject assessments for other images

4. プライバシー保護の客観評価指標

プライバシー保護の客観的評価指標である秘匿性を計算するため、第3章で得られた主観評価指標と一致する客観指標を検討した。Fig. 4 に示した主観評価の結果より、画質が悪いほど主観評価スコアは高い傾向がある。そこで、秘匿性の客観評価指標として、画像の画質評価指標が有用であると考え、3つの画質指標を比較、検討した。

4.1 比較・検討した評価指標

本稿では Mean Squared Error (MSE), Structure Similarity (SSIM), Most Apparent Distortion (MAD) の3つの画質評価指標の、秘匿性指標としての有効性を検討する。これらの画質評価指標は、通常画像 (Fig. 3a) と評価画像 (Fig. 3b) の誤差を評価する。なお、画像の各画素の RGB 情報はすべて、

$$L = 0.2989R + 0.5870G + 0.1140B \quad (1)$$

によって輝度値 L に変換してから各評価値を計算した。

まず、MSE は次式によって計算される。

$$MSE = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N (I(m,n) - K(m,n))^2 \quad (2)$$

ここで、 $I(m,n)$ は評価画像の m 行 n 列目の画素の輝度値、 $K(m,n)$ は通常画像の m 行 n 列目の画素の輝度値を表しており、評価画像および通常画像のサイズを $M \times N$ pixels とした。この指標は、評価画像 I と通常画像 K の画素ごとの差の平均二乗和である。MSE は 0 以上の値を取り、値が大きいほど通常画像との差が大きいため、秘匿性は高くなると推測される。

次に、SSIM は、一般的に画像を小領域に分けて計算される [2]。画像 I 、 K の p 番目の小領域をそれぞれ i 、 k とすると、SSIM は、

$$SSIM(p) = \frac{(2\mu_i\mu_k + C_1)(2\sigma_{ik} + C_2)}{(\mu_i^2 + \mu_k^2 + C_1)(\sigma_i^2 + \sigma_k^2 + C_2)} \quad (3)$$

と表される。ここで、 μ_i 、 μ_k は小領域 i 、 k の輝度の平均値、 σ_i^2 、 σ_k^2 は小領域 i 、 k の輝度の分散、 σ_{ik} は共分散を表している。また、参考文献 [2] に従い、 $C_1 = 6.5025$ 、 $C_2 = 58.5225$ とした。本稿では、1つの評価画像に対して1つの評価値を得るために、小領域ごとに計算された $SSIM(p)$ の平均値を当該評価画像の $SSIM$ 値とした。この指標は、分散を利用することで画像のコントラストの違い、共分散を利用することで画像のエッジ構造の違いも加味しており、MSE に比べ高次元特徴を利用した評価指標である。SSIM は 0 から 1 の値を取り、値が低いほど通常画像との差が大きいため、秘匿性が高くなると推測される。

最後に、MAD は、人間は画質評価に2つの戦略を使い分けているという仮説のもとに提案された指標である [3]。1つ目の戦略は、比較的画質の良い画像を見る場合、画像の中から画質劣化の原因となるノイズやボケを探し出す detection-based 戦略である。2つ目の戦略は、画質の悪い画像を見る場合、ノイズの中から被写体を探し出す appearance-based 戦略である。MAD は、評価画像を detection-based 戦略による評価値 d_{detect} と、appearance-based 戦略による評価値 d_{appear} を用いて、

$$MAD = (d_{detect})^\alpha (d_{appear})^{1-\alpha} \quad (4)$$

$$\alpha = \frac{1}{1 + \beta_1 (d_{detect})^{\beta_2}} \quad (5)$$

で計算される。 β_1 、 β_2 はフリーパラメータであり、本項では、参考文献 [3] で算出された値 $\beta_1 = 0.469$ 、 $\beta_2 = 0.130$ を使用した。 d_{detect} および d_{appear} の計算方法は参考文献 [3] を参照されたい。 d_{detect} の計算過程では、画像の輝度値と人間が感じる明るさの非線形性や、人間の視覚系が持つコントラスト感度も加味されている。したがって、MAD は、

MSE や SSIM よりも人間の主観に近い画質評価を可能とすることが示されている。MAD は 0 以上の値を取り、値が大きいほど通常画像との差が大きく、秘匿性が高くなると推測される。

4.2 評価範囲の走査

MPC で撮影された画像は、撮影位置が異なる多重像になるため、評価画像の平行移動には対応しない上記評価指標を導入することは困難である。この課題に対応するため、評価する範囲を走査し、その極値を評価値として利用する。本稿では、通常画像の顔領域 (Fig. 3a に示した 63×44 pix の範囲) を切り抜き、評価画像上を走査しながらそれぞれの個所で評価値を計算し、その極値を、この画像の評価値とした。

5. シミュレーション結果

Fig. 3b に示した 13 枚の評価画像について、3 つの客観評価値を求めた結果を Fig. 6 に示す。それぞれの評価指標ごとに、通常画像との誤差が小さいものから順に左から右へ並べている。また、各評価画像の上部には Fig. 4 と同様の生成パターンのインデックスを示し、下部にはそれぞれの評価値を示す。

Fig. 6 より、どの評価指標についても、通常画像との誤差が小さいほど顔の輪郭や目・鼻の位置が正確に識別でき、通常画像との誤差が大きいほど識別は困難となっており、これは Fig. 4 に示した主観評価と類似する結果である。特に、主観評価によって最も人物を特定しやすいと判定された評価画像 A~D は、MSE, SSIM, MAD のいずれの評価指標によっても最も通常画像との誤差が小さいと判定された。以上のことから、MPC で撮影された画像や復元画像の秘匿性を評価する客観指標として、画質評価指標が有用であることがわかった。

一方で、MPC にノイズを付与した画像を SelfDeblur 法によって復元した画像である評価画像 F, G, K の順位は、評価指標ごとにばらつきがあり、特に SSIM による評価順位は主観評価よりも明らかに低かった。これらの画像では、通常画像と評価画像の間の共分散が大きく増加するため、共分散を利用して決定される SSIM の評価値は低い傾向を示すと推測される。

さらに、主観評価とそれぞれの客観評価の結果が一致しているか確認するために、横軸を各評価値、縦軸を主観評価値としたグラフを Fig. 7 に示す。各点は 1 つの評価画像を表しており、主観評価値は平均±標準偏差のエラーバーも表示している。MSE および MAD の評価値が増加するほど主観評価値が増加し、SSIM の評価値が増加するほど主観評価値も減少する傾向があり、これらの客観指標は主観

評価と相関を持っていることがわかる。しかし Fig. 7b より、SSIM が小さい側から 3~5 番目の点 (順に画像 F, G, I に対応) は、SSIM の値が増加するほど主観評価値が増加している。また、SSIM 値が約 0.45 の点 (画像 H に対応) も、その前後の点よりも主観評価は明らかに高い値を示した。このように、評価順位の逆転が頻繁に生じていることから、SSIM による秘匿性評価は、主観評価と一致していない。一方で、MSE や MAD による評価順位は主観評価との逆転が少なく、逆転してもエラーバーの高さが重なっており有意な差であるとは言えないため、これらの客観指標は主観評価とよく一致している。

今回比較している 3 つの客観指標について、計算時間も確認した。Intel(R) Core(TM) i9-9980HK CPU @ 2.40GHz において、Fig. 3 に示した評価画像 (解像度 100×100 pix) 1 枚あたりの評価値の計算時間は、MSE は 0.0360 ± 0.0036 秒、SSIM は 3.69 ± 0.07 秒、MAD は 15.1 ± 0.1 秒であった。つまり、MSE 値は MAD 値の約 420 倍高速に計算できる。前述の通り、MSE と MAD は同程度に主観評価とよく一致する結果を示すが、計算時間は MSE の方が圧倒的に速いことから、今回比較した 3 つの客観指標のうち、MSE が最も実用的な秘匿性評価指標である。

6. まとめ

本稿では、MPC で撮影した画像を対象として、プライバシー保護性能を評価するための客観指標として利用可能な評価基準について検討を行った。

プライバシー保護性能とは「主観的な、画像からの人物特定容易性」であると仮定し、その客観的な評価指標を求めるために、画質評価指標である、MSE, SSIM, MAD を評価した。画像からの人物特定容易性の主観評価値と、各評価指標の関係性を調べることにより、MSE と MAD は同等に主観評価と一致していたが、SSIM 復元画像に関して主観評価とは一致しない結果を示した。

また、処理時間を比較し、MSE が他の手法に比べて高速であることがわかった。

このことから、MSE は、主観評価との一致度が高く、計算時間も短いため、今回の目的に対して最も実用的な指標であることがわかった。

今後は、人物特定に限定しない他の画像の秘匿性評価にも MSE が有効であるか検討する。

参考文献

- [1] Y. Ishii, S. Sato and T. Yamashita, "Privacy-Aware Face Recognition with Lensless Multi-Pinhole Camera," in *European Conference on Computer Vision Workshops*, 2020.
- [2] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image

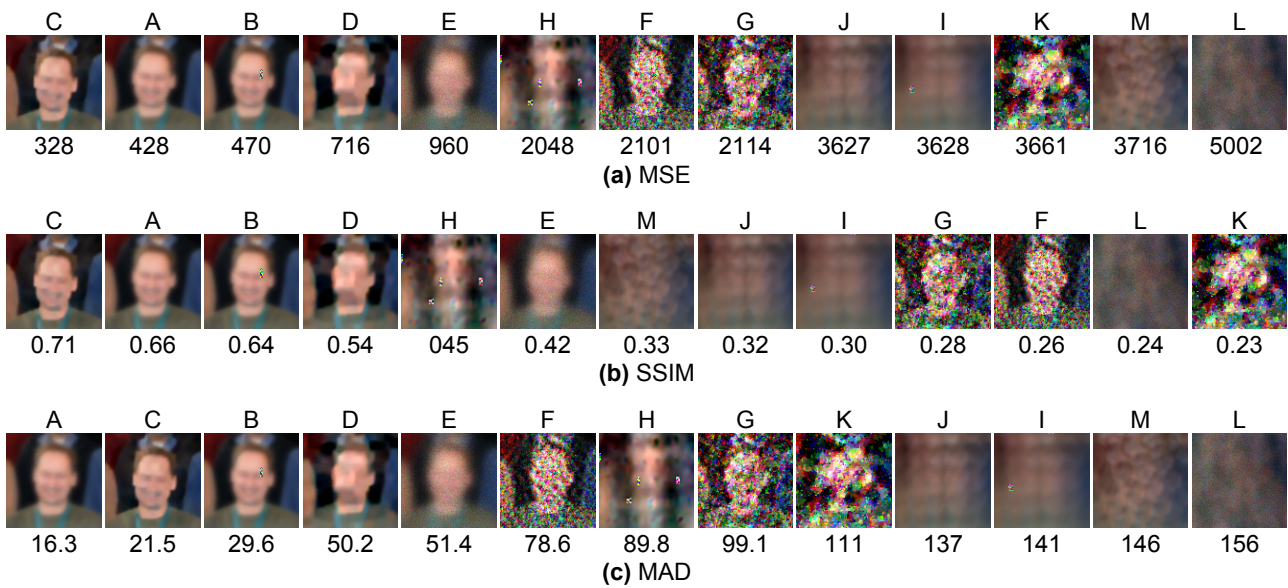


Figure 6. Results of objective assessment

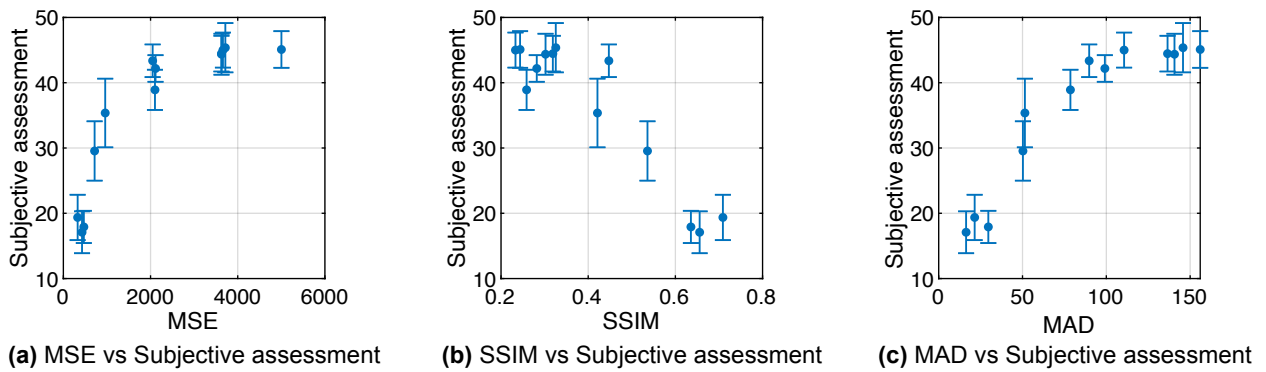


Figure 7. Objective assessments vs Subjective assessment

Quality Assessment: From Error Visibility to Structural Similarity,"
Transactions on Image Processing, vol. 13, no. 4, pp. 600-612,
 2004.

- [3] E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *J Electron Imaging*, vol. 19, no. 1, 2010.
- [4] S. Shao, Z. Zhao, B. Li, T. Xiao, G. Yu, X. Zhang and J. Sun, "CrowdHuman: A Benchmark for Detecting Human in a Crowd," in *arXiv preprint arXiv:1805.00123*, 2018.
- [5] D. Ren, K. Zhang, Q. Wang, Q. Hu and W. Zou, "Neural Blind Deconvolution Using Deep Priors," in *Computer Vision and Pattern Recognition*, 2020.