

# 考慮視覺敏感度之影像品質評估

郭天穎、徐詩媛、謝昌利  
國立台北科技大學電機工程系

**摘要** — 數位化的影像傳輸或壓縮會對影像造成失真的問題，可經影像品質評估來比較差異，而最終接收端為人眼，所以現今的評估方式會更考慮人眼視覺敏感度。本文所提出的評估方法是改進視覺信號保真度(VIF)，使用 Haar 小波轉換、對數-賈伯濾波器、對比敏感度係數(CSF)、Laplacian filter、SSIM 做評估，使得影像評估著重於視覺敏感的特性和影像物件的部份，經實驗測試證明之後發現，本文的評估效能和時間複雜度超越目前最新的 Perceptual Quality Metric With Internal Generative Mechanism(PQM)。<sup>1</sup>

## 一、簡介

數位化影像的取樣和處理過程，常造成影像的失真或雜訊的影響，品質評估是為了評估失真過後的影像與原本影像的差異量，而最終接收端為人眼，所以不僅僅是考慮兩者的誤差量，更需要考慮失真影像對於人眼的感知程度的影響。目前常見的影像評估方式有 SSIM、VIF、IW-SSIM、FSIM 等...，最新的為 Perceptual Quality Metric PQM[4]，這些演算法都有將人眼的視覺影像考慮進去，不過最先考慮人眼視覺的影響的評估方法是 Structural similarity (SSIM)[2]和 Visual information fidelity (VIF)[3]的這兩個影像評估演算法，SSIM 將影像分成亮度、對比、結構，使用局部的計算三者的相似性，合成起來最後作平均，得到貼近人眼的評估，即為 SSIM 的評估演算法；VIF 使用金字塔小波之後的影像做互信息(Mutual information)，最後將所有的互信息總和做比值，得到影像評估結果。

本論文所提出的方法是基於 VIF 做改進，首先將參考影像和失真影像作 Haar 小波轉換，濾除掉變化量對於視覺較不敏感的高頻成分，保留影像變化對於人眼視覺重要的低頻成分，接著經過對數-賈伯濾波器[7]將兩張保留低頻成分影像的頻譜分解為多個尺度的水平和垂直方向頻帶，替代 VIF 原本的金字塔小波轉換取不同頻帶與方向的效果，根據每個頻帶在空間域的響應，擷取失真影像上人眼較敏感的特徵和相對應的參考影像特徵，再利用 VIF 的方法計算兩張影像局部的互信息，估測影像局部經人類視覺系統通道傳送到大腦的信息量，最後使用頻譜 Laplacian filter 和 SSIM 偵測參考影像上人眼較敏感的物件區域和失真影像，以此作為局部信息合成權重的依據，累加兩張影像所有局部信息後再求比值，即可得到影像評估結果。

為了將影像更符合人眼所觀看的現象，我們使用了對比敏感函數(CSF)將對數-賈伯濾波器的尺度所分解的影像給予權重，且能依照主觀環境條件選擇人眼較為敏

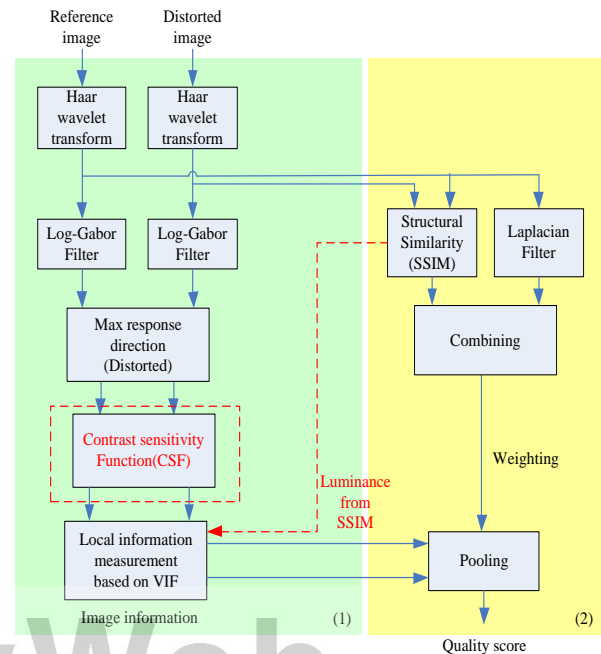
感的頻帶特徵；由於 VIF 本身著重於信息量的比較屬於紋理失真的偵測，無法有效地評估亮度(Luminance)失真，所以我們並在計算影像信息量的時候，同時使用 SSIM 評估參考影像和失真影像的亮度相似性，以彌補 VIF 對於平滑失真的缺失。

經影像資料庫測試結果發現，本文所提出的 WSS-VIF 相較於原本的 VIF[3]，影像評估的準確性和一致性都有很顯著的提升，運算複雜度也比原本的 VIF 低，影像評估的效能和效率更是超越近年最新提出的 PQM[4]、IW-SSIM[5]和 FSIM[6]。

本論文的架構如下所示，第二章詳細說明本文所提出的 WSS-VIF 影像評估方法之理論、架構和實作流程，第三章是實驗結果，最後為本論文之結論。

## 二、WSS-VIF

本文提出一個基於 VIF[3]作改進的影像評估方法，稱為敏感度加權的視覺信號敏感保真度，簡稱 WSS-VIF。VIF 影像評估大致可分為兩個步驟，第一個步驟為計算參考影像和失真影像的互信息；第二個步驟是將兩張影像所有的互信息累加求比值，合成整體評估結果。我們同時對這兩個步驟進行改進，WSS-VIF 影像評估之流程如圖一所示。



圖一：WSS-VIF 影像評估流程

<sup>1</sup> 本研究由國科會贊助，計畫編號 NSC 101-2221-E-027-789-MY2。

## 2.1 計算影像的互信息

WSS-VIF 的第一個步驟是將參考影像和失真影像作 Haar 小波轉換，保留包含影像重要資訊的低通影像，接著使用對數-賈伯濾波器[7]將兩張低通影像的頻譜分解為多個尺度的水平和垂直方向頻帶，根據每個頻帶在空間域的響應，擷取失真影像上人眼較敏感的特徵和相對應的參考影像特徵，再利用 VIF[3]的方法計算兩張影像局部的互信息，估測影像局部經人類視覺系統通道傳送到大腦的信息量。

### 2.1.1 Haar 小波轉換濾除高頻成分

低通濾波器、高通濾波器和次取樣濾波器組成。影像經過一階 Haar 小波轉換後，可分解為一張低通影像 LL，還有三張高通影像，分別為 HL、LH 和 HH。低通影像內含最多影像資訊，高通影像則表示原始影像上像素值有變化的地方，在這邊捨棄掉人眼較不敏感的高通影像，留下低通影像，讓後續的影像評估著重於影像重要的資訊。

### 2.1.2 對數-賈伯濾波器優先擷取失真影像特徵

對數-賈伯濾波器[7]在頻域上的轉換函數為  $G(\omega, \theta)$ ，如(1)式，第一項可決定濾波器的半徑， $\sigma_r$  為濾波器的頻寬， $\omega_0$  為濾波器中心頻率，可透過調整中心頻率  $\omega_0$  產生不同尺度的濾波器；第二項可決定濾波器的角度， $\sigma_\theta$  為濾波器的角頻寬， $\theta_j$  為濾波器的方向角，可透過調整方向角  $\theta_j$  產生不同方向的濾波器。基於人眼對於影像水平和垂直方向的特徵較敏感，本文使用 4 個尺度、每個尺度有 2 個方向的對數-賈伯濾波器，將影像頻譜分解為 4 個尺度的水平和垂直方向頻帶，以此偵測空間域上不同尺度的水平和垂直方向特徵。

$$G(\omega, \theta) = \exp\left(-\frac{\left(\log \frac{\omega}{\omega_0}\right)^2}{2(\log \sigma_r)^2}\right) \exp\left(-\frac{(\theta - \theta_j)^2}{2\sigma_\theta^2}\right) \quad (1)$$

一般在作主觀影像評估時，人眼會先觀察到失真影像上較明顯的特徵，再與參考影像的相同位置做比較，判別影像的失真程度，所以在作影像評估的時候，應該優先考慮失真影像上較明顯的特徵。根據上述的原因，首先我們使用對數-賈伯濾波器將失真影像頻譜分解為 4 個尺度的水平和垂直方向頻帶，比較同一個尺度的水平和垂直方向頻帶在空間域的響應，取相同位置上較大的響應作為該像素點在該尺度的響應；接著再使用對數-賈伯濾波器將參考影像頻譜也分解為 4 個尺度的水平和垂直方向頻帶，在同一個尺度相同位置的水平和垂直方向頻帶響應中，取跟失真影像較大響應相同方向的響應作為該像素點在該尺度的響應，本文透過此方法將失真影像和參考影像 4 個尺度的水平和垂直方向頻帶響應合成 4 個尺度的最大頻帶響應和相對應頻帶響應。

### 2.1.3 對比敏感度函數(CSF)

人眼視覺系統(HVS)研究提供了數學模型表示人類是如何看周圍的世界，模型已經發現人眼的視覺在亮度

及顏色的特徵。最初由 Mannos 和 Sakrison 提出關於亮度(或灰階)的影像 CSF 模型，如公式(2)所示。

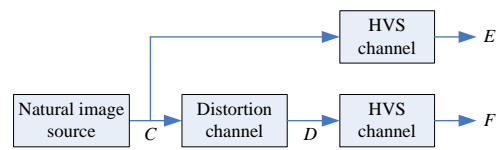
$$H(f) = 2.6(0.192 + 0.114f)e^{[-(0.114f)^{1.1}]} \quad (2)$$

空間頻率  $f$  單位為(cycles/degree)。CSF 函數曲線對於對比度的刺激很明顯，但是不同於亮度的刺激，人類的視覺特性顯示為一種帶通濾波器，對於對比度刺激的敏感度，在整個空間頻率上是相當一致，也就是說該帶通濾波器的範圍內算出的對比敏感度的係數是一樣的權重。

選擇適當的對數-賈伯濾波器頻帶，使得可以和人眼所察覺得敏感度更加接近。

### 2.1.4 VIF 計算影像局部互信息

VIF[3]使用信息理論的概念來做影像評估，模擬影像信號經過人類視覺系統通道，由大腦擷取感知信息。它主要的架構如圖一所示，首先使用小波轉換將參考影像分解為多尺度的水平和垂直方向頻帶，利用小波域的高斯混合尺度模型(Gaussian scale mixtures model, GSM)將參考影像的每個頻帶轉成高斯向量隨機場。接著再使用小波轉換將失真影像也分解為多尺度的水平和垂直方向頻帶，透過參考影像頻帶和相對應失真影像頻帶計算失真通道(Distortion channel)，失真通道可利用小波域上信號的衰減和雜訊來模擬影像的失真。假設參考影像頻帶轉成高斯向量隨機場後為  $C$ ，可透過  $C$  與失真通道合成失真影像頻帶  $D$ ， $E$  和  $F$  則分別為  $C$  和  $D$  經過人類視覺系統通道(HVS channel)後輸出至大腦的結果。



圖二：VIF 的架構

最後計算  $C$  和  $E$  的互信息  $I(C; E)$ ，估測參考影像經人類視覺系統通道傳送到大腦的信息量；再計算  $C$  和  $F$  的互信息  $I(C; F)$ ，估測相對於參考影像，失真影像經人類視覺系統通道傳送到大腦的信息量，將兩張影像所有頻帶的互信息累加求比值，即可得到 VIF 影像評估值，如(3)式，其中  $j$  為影像被分解的頻帶數量，高斯向量隨機場條件用符號“ $|S^j$ ”代表。

$$VIF = \frac{\sum_{j \in \text{subbands}} I(C^j; F^j | S^j)}{\sum_{j \in \text{subbands}} I(C^j; E^j | S^j)} \quad (3)$$

在這裡我們將 VIF 的小波與上的水平與失真的特徵影像替換成我們影像經 Haar 小波轉換和對數-賈伯濾波器求最大響應的特徵影像。

## 2.2 合成整體評估結果

第二個步驟同時使用 SSIM[2] 和 Laplacian filter 偵測視覺敏感區域，以此產生每個頻帶響應局部信息合成的權重，累加兩張影像所有頻帶響應的信息後再求比值，即可得到整體的影像評估結果。

### 2.2.1 合成整體評估結果

首先使用 SSIM 計算兩張影像的評估特徵圖，因為 SSIM 其評估值越高代表局部的影像品質越好，基於人眼對於影像上失真較嚴重的區域較敏感，可用反相評估特徵圖，突顯影像上失真較嚴重的區域，方便後續的處理。反相評估特徵圖分別為失真較嚴重和失真不嚴重區域，產生失真特徵圖，也另外將亮度相似性提出，使用在權重的總合計算的時候，局部對照失真互信息量，以彌補 VIF 亮度不足的部分，由於人眼視覺對於亮度的敏感度比較低，所以將亮度相似性的值開根號再帶入。

### 2.2.2 Laplacian filter 方法偵測物件區域

接著使用 Laplacian filter 和高斯平滑濾波器求得參考影像的邊緣顯著圖，通常影像邊緣和紋理較多的區域可能包含重要的物件，也是人眼較敏感的地方，我們使用邊緣顯著圖像素值平均作為門檻值，擷取邊緣顯著圖上包含物件的區域，產生物件特徵圖，分別標示將包含物件的區域和不包含物件的區域。

### 2.2.3 影像評估結果

將失真特徵圖和物件特徵圖作結合，並給予適當的影像評估權重，當影像局部同時為失真較嚴重和物件區域時，將權重設為最高；當影像局部為失真較嚴重區域但非物件區域時，將權重中間；當影像局部為物件區域但非失真較嚴重區域時，將權重中間；其餘非失真較嚴重且非物件區域敏感度比較低，則將權重設為最小，產生影像評估合成權重圖  $W$ 。

$$WSS-VIF = \frac{\sum_{j,N} I(c^{N,j}; f^{N,j}|s^{N,j}) \cdot L(N)^{0.5} \cdot W(N)}{\sum_{j,N} I(c^{N,j}; e^{N,j}|s^{N,j}) \cdot W(N)} \quad (6)$$

## 三、實驗結果

### 3.1 影像資料庫與評估效能指標

TID2008[9]和 CSIQ[10]為目前文獻上最大的兩個影像資料庫，TID2008 影像資料庫總共有 1700 張測試影像，內含 17 種影像失真類型；CSIQ 影像資料庫總共有 866 張測試影像，內含 6 種影像失真類型。

通常評估主觀分數與客觀分數的關係會使用四種常見的效能指標來比較影像評估方法的效能，其中兩種為 Spearman 等級相關係數(SRCC)和 Kendall 等級相關係數(KRCC)，它們只考慮主觀分數與客觀分數排序的差異性，可用來評估影像評估結果與主觀分數的一致性，另外兩種為 Pearson 線性相關係數(PLCC)和均方根誤差(RMSE)，它們可用來評估影像評估方法的準確性，因為每種主觀分數和客觀分數的值域都不太一樣，為了方便比較，可先做迴歸分析，找出客觀分數與主觀分數的非線性映射關係，再利用主觀分數和非線性映射後的客觀分數計算 PLCC 和 RMSE，可透過迴歸分析找出。一個好的影像評估演算法必須具備較高的 SRCC、KRCC 和 PLCC，還有較低的 RMSE。

### 3.2 VIF 改進前後效能比較

本文提出的方法是基於 VIF[3]作改進，為了證明我們

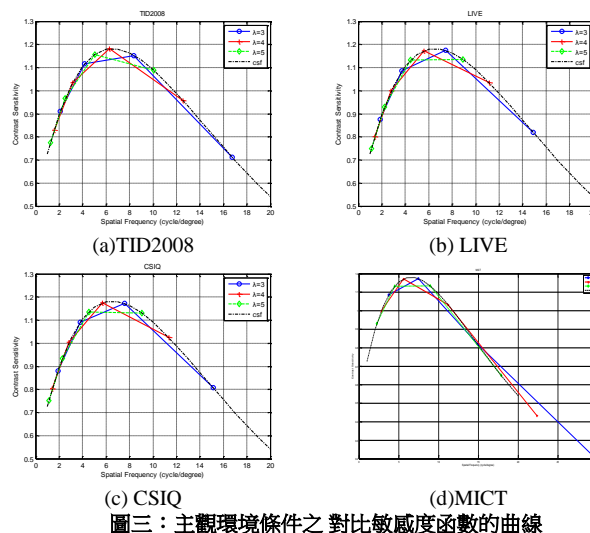
加入的視覺敏感特性有助於提升影像評估的效能，首先測試 VIF 改進前後評估效能的差異，TID2008[9]影像資料庫測試結果如表 I 所示，我們分別測試原本的 VIF、加入 Haar 小波轉換濾除影像高頻成分的 VIF、加入對數-賈伯濾波器[7]優先考慮失真影像特徵的 VIF、加入 Laplacian filter 偵測影像物件區域的 VIF、加入 SSIM 偵測失真區域的 VIF 和本文提出的 WSS-VIF，經測試結果發現濾除影像高頻成分、優先考慮失真影像特徵和偵測影像物件區域都可以有效的提升 VIF 影像評估的效能，讓影像評估的結果更貼近人眼察覺的影像品質。而本文所提出的 WSS-VIF 因為同時使用這三種視覺敏感特性，所以相較於原本的 VIF，影像評估準確性和一致性都有非常顯著的提升。

表 I  
VIF 改進前後效能比較

Model	SRCC	KRCC	PLCC	RMSE
VIF	0.7491	0.5860	0.8084	0.7899
Haar transform	0.8461	0.6685	0.8624	0.6794
Log-Gabor filter	0.8387	0.6678	0.8699	0.6620
Structural similarity	0.7840	0.6080	0.8316	0.7453
Laplacian filter	0.7744	0.6039	0.8239	0.7605
SS-VIF	0.8975	0.7287	0.9050	0.5709
SS-VIF+Luminance	0.8959	0.8929	0.9098	0.8487
SS-VIF+CSF	0.9021	0.7341	0.9113	0.5525
WSS-VIF	<b>0.9014</b>	<b>0.7337</b>	<b>0.9009</b>	<b>0.5538</b>

### 3.3 主觀環境條件效能比較

使用主觀環境條件下，去選取對數-賈伯濾波器的週期，週期與中心頻率有關係，調整週期可以調整尺度範圍，以目前有提供的主觀環境條件所計算出最符合對比敏感度函數的曲線。



圖三：主觀環境條件之對比敏感度函數的曲線

### 3.4 與常見的方法效能比較

使用常見的影像評估方法 PSNR、SSIM[2]、MS-SSIM[11]、VIF[3]、PQM[4]、IW-SSIM[5]和 FSIM[6]與 WSS-VIF 做比較，TID2008[9]和 CSIQ[10]影像資料庫的測試比較結果如表 II 和表 III 所示，其中粗體字代表每一個評估指標中效能最好的分數，經測試結果發現，我們提出的 WSS-VIF 此中都具有最好的一致性和準確性。

擁有最高的 SRCC、KRCC 和 PLCC，還有較低的 RMSE 的評估值。

表 II  
WSS-VIF 與常見影像評估方法在 TID2008 效能比較

Model	SRCC	KRCC	PLCC	RMSE
PSNR	0.5531	0.4027	0.5734	1.0994
SSIM	0.7749	0.5768	0.7732	0.8511
MS-SSIM	0.8542	0.6568	0.8451	0.7173
VIF	0.7491	0.586	0.8084	0.7899
IW-SSIM	0.8559	0.6636	0.8579	0.6895
FSIM	0.8805	0.6946	0.8738	0.6525
PQM	0.8945	0.7226	0.9044	0.5725
WSS-VIF	<b>0.9014</b>	<b>0.7337</b>	<b>0.9009</b>	<b>0.5538</b>

表 III  
WSS-VIF 與常見影像評估方法在 CSIQ 效能比較

Model	SRCC	KRCC	PLCC	RMSE
PSNR	0.8058	0.6084	0.8	0.1575
SSIM	0.8756	0.6907	0.8613	0.1334
MS-SSIM	0.9133	0.7393	0.8991	0.1149
VIF	0.9195	0.7537	0.9277	0.098
IW-SSIM	0.9213	0.7529	0.9144	0.1063
FSIM	0.9242	0.7567	0.912	0.1077
PQM	0.9403	0.7881	0.9281	0.0977
WSS-VIF	<b>0.9483</b>	<b>0.8011</b>	<b>0.9477</b>	<b>0.0838</b>

### 3.5 影像評估散點圖

圖三為 VIF[3]、IW-SSIM[5]、FSIM[6]和 WSS-VIF 在 TID2008[9]影像資料庫測試結果與主觀分數的散點圖和非線性映射曲線，其中 X 軸為影像評估的客觀分數，Y 軸為影像資料庫的主觀分數。觀察散點圖和非線性映射曲線可證實，WSS-VIF 影像評估的結果與主觀分數具有比較好的線性關係，整體的相關性和一致性明顯比近年來最新的方法好。

### 3.6 運算複雜度

在運算複雜度比較是使用電腦硬體配備為 Intel(R) Core(TM)2 Duo CPU E7300 @2.66GHz, 3.50GB RAM，並以 Matlab R2010a 軟體作模擬和測試。表 WSS-VIF 與其他常見的影像評估方法平均測試一張 512×512 影像所需的時間，測試方式是使用 IVC 資料庫的影像，共 185 張影像，每張影像使用客觀評估演算法評估 10 次，總共 1850 次所得的平均時間，如表 IV 所示。

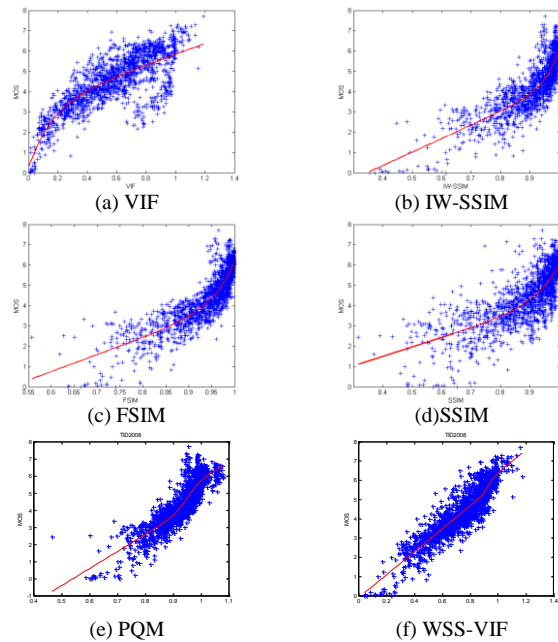
表 IV  
運算複雜度比較(秒/影像)

Model	Time	Model	Time
IW-SSIM	1.1098	PSNR	0.0027
FSIM	0.8403	SSIM	0.0605
PQM	30.5951	MS-SSIM	0.1566
WSS-VIF	<b>0.6786</b>	VIF	1.8694

### 結論

本文提出一個基於 VIF[3]作改進的影像評估方法，稱為 WSS-VIF，參考了主觀評估時的環境因素，進而使用了人眼視覺的特性，使得評估值能更符合人眼所觀察，並提升了與主觀分數一致性。偵測紋理失真的部份，是依照當時主觀的偵測狀況模擬當時的可觀察的像素，並用對比敏感度函數(CSF)找出敏感的空間頻率的權重，使得敏感度係數得到對數-賈伯濾波器的頻帶給予不同權重部分加強頻帶的敏感度，然後帶入 VIF 計算互信息量。再提出平滑失真的亮度失真的補償了計算失真度的互信息量，接著再用強調物體區域及失真區域的權重，合成整張影像的評估值。

在實驗結果以本文提出改進方法作比較，WSS-VIF



圖三：影像評估散點圖

影像評估加入了許多視覺敏感特性，其中對比度敏感函數(CSF)和亮度相似性，使得 WSS-VIF 相較於 VIF 和普遍常見的影像評估方式的整體評估效能高和實驗複雜度低，證明了 WSS-VIF 能更進一步提升影像評估的整體效能和一致性。以本文所提出的方法，經 TID2008[9]和 CSIQ[10]影像資料庫測試結果相較於原本的 VIF，WSS-VIF 影像評估的效能和效率已經超越目前最新且效能最好的方法，運算複雜度也比原本的 VIF 低，影像評估的效能和效率更是超越近年最新的方法。

### 參考文獻

- [1] Z. Wang and A. C. Bovik, *Modern Image Quality Assessment*. San Rafael, CA: Morgan & Claypool Publishers, Mar. 2006.
- [2] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600-612, Apr. 2004.
- [3] H.R. Sheikh and A.C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430-444, Feb. 2006.
- [4] Jinjian Wu, Weisi Lin, Guangming Shi, and Anmin Liu, "Perceptual Quality Metric With Internal Generative Mechanism," *IEEE Trans. Image Process.* vol. 22, no. 1, pp. 43-54, Jun. 2013.
- [5] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1185-1198, May 2011.
- [6] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378-2386, Aug. 2011.
- [7] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *J. Opt. Soc. Amer. A*, vol. 4, no. 12, pp. 2379-2394, Dec. 1987.
- [8] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," *IEEE Conf. CVPR '07*, pp. 1-8, Jun. 2007.
- [9] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "TID2008 - A database for evaluation of full-reference visual quality assessment metrics," *Adv. Modern Radioelectron.*, vol. 10, pp. 30-45, 2009.
- [10] E. C. Larson and D. M. Chandler, *Categorical Image Quality (CSIQ) Database 2009* [Online]. Available: <http://vision.okstate.edu/csiq>
- [11] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," in *Proc. IEEE Asilomar Conf. Signals, Syst. Comput.*, pp. 1398-1402, Nov. 2003.