# Illumination-Invariance and Nonlocal Means Based Super Resolution

Mengyan Wang, Jiaying Liu*, Wei Bai and Zongming Guo
Institute of Computer Science and Technology, Peking University, Beijing, P. R. China, 100871

***Abstract*—In this paper, we propose a novel algorithm for multi-frame super resolution (SR) with illumination-invariance. Traditional multi-frame SR methods fail to handle images with illumination changes, so in our approach, we adjust the contrast between different search windows and select proper candidate patches to take full advantage of intensity information. We simplify Speed Up Robust Features to get local structure information and incorporate the local structure information into similarity measurement, which does not change significantly in complex illumination situation. By combining intensity and structure information in a proper way, our algorithm *Illumination-Invariant Nonlocal Means SR* could find more potential similar patches in frames where there are illumination changes than Nonlocal Means SR (NLM SR). Experimental results demonstrate that our algorithm has better performance both in objective and subjective perception with complex illumination conditions and is comparable to NLM SR in stable illumination situation.**

## I. Introduction

Super resolution is a long-standing research area in the image processing community. Multi-frame super resolution is a traditional super resolution method, aiming at recovering a high-resolution (HR) image from a series of low-resolution (LR) images. By using redundant information in LR frames, we can reconstruct a high quality HR image with less aliasing and fewer artifacts. The key to conventional multi-frame super resolution algorithm is to know the sub-pixel displacements between different frames. But since it is very difficult to acquire precise motion estimation, conventional multi-frame super resolution algorithms are severely restricted. To solve this problem, Potter *et al.* [1] proposed NLM super resolution algorithm with no explicit motion estimation. By replacing every pixel with a weighted average of its neighborhood, NLM SR could get better HR image.

On the basis of NLM SR, various improvements were proposed to make similar patches stand out and give them proper weights. Cheng *et al.* [7] proposed a SR approach using a mobile search strategy. By adjusting search window size and block size adaptively, this algorithm could get better visual quality. Our previous work [4] used the search window relocation and rotation-invariance measurement to improve NLM SR.

All the aforementioned methods used intensity information, *i.e.*, the value of a pixel, as a way to measure the similarity of patches and they assumed the illumination condition was stable in images. However, natural images can be easily affected by illumination. Thus, when there are illumination changes between adjacent frames, those methods can not get good performance due to the loss of similar patches or the inaccuracy of the weight of patches.

Considering illumination changes problem, our previous work [6] proposed a novel method and applied it into denoising. We adjust the contrast between adjacent frames. So for the areas in an image which have same reflection to illumination, our previous work could eliminate the influence of illumination. But since different parts of an image have different reflections under complex illumination conditions, this method was not good enough in boundaries between foreground and background due to the inappropriate selection of candidate patches.

Taking account of the above issues, we propose an illumination-invariance measurement to calculate the similarity between different patches. Instead of just adjusting contrast between frames, we also select proper candidate patches, which guarantees us to find more similar patches. Due to the relative stability of local structure information in complex illumination conditions, we also take the structure information of a patch into account. We simplify Speed Up Robust Features (SURF) [3] to get local structure information descriptor. Combining structure information and modified intensity information together and giving them suitable weights, we can measure more accurate similarity of patches even in complex illumination conditions.

The rest of this paper is organized as follows. In Section II, we introduce NLM SR and discuss its drawbacks. Section III focuses on our proposed Illumination-Invariance NLM SR algorithm. Experimental results are shown and analyzed in Section IV. A brief conclusion is given in Section V.

## II. Review on Nonlocal Means Super Resolution

Traditional NLM used redundant information in similar patches for denoising. Then, Elad *et al.* extended this idea to super resolution in [1] and his method is called NLM SR.

$Y_t$ and $Y_{t'}$ are original $t$-th and $t^{'}$-th frames in a HR sequence which is interpolated from the LR sequence. For a reference patch centered at $(k,l)$ in $Y_t$, candidate patches are selected from a predetermined search window in frames around $Y_t$. The similarity between the reference patch and a candidate patch centered at $(i,j)$ in $Y_{t'}$ is calculated as follow:

$$w(k,l,i,j) = \frac{1}{C(k,l)} exp\left\{-\frac{\|R_{k,l}Y_t - R_{i,j}Y_{t'}\|_2^2}{2\sigma^2}\right\}, \quad (1)$$

where $R_{i,j}$ is an operator which extracts a patch with a size $(q*q)$ from an image and get a vector whose length is $q^2$.

*Correspongding author

This work was supported by National Natural Science Foundation of China under contract No. 61101078, National Basic Research Program(973 Program) of China under contract No.2009CB320907 and Doctoral Fund of Ministry of Education of China under contract No. 20090001120027.

$C(k,l)$ is the normalization constant. $\sigma$ controls the intensity difference between two patches.

After assigning proper weights to different patches, we can estimate the value of pixel $(k,l)$ as follows:

$$V(k,l) = \frac{\Sigma_{t' \in [1,\cdots,T]} \Sigma_{(i,j)\in N(k,l)} w(k,l,i,j) Y_{t'}(i,j)}{\Sigma_{t' \in [1,\cdots,T]} \Sigma_{(i,j)\in N(k,l)} w(k,l,i,j)}, \quad (2)$$

where $N(k,l)$ means the search window around pixel $(k,l)$.

Since NLM SR uses intensity information to calculate the similarity of two patches, it works well when the illumination condition is stable in the given sequence. But because the gray-level of a pixel can be easily affected by illumination, NLM SR can not find similar patches and is not suitable in complex illumination condition. To solve this problem, we propose illumination-invariant algorithm to improve NLM SR by measuring the similarity between different patches properly.

## III. Illumination-Invariance Nonlocal Means SR

We propose an illumination-invariance method to calculate similarity between two patches. First, we consider intensity similarity and structure similarity separately in Section III-A and Section III-B . Then, we combine these two methods together properly to get our illumination-invariant similarity measurement in Section III-C. Finally, our algorithm reconstructs high resolution frames as NLM SR does.

### A. Intensity Similarity Measurement

In this section, we focus on measuring intensity similarity between different patches. NLM SR uses SSD (sum of squared error) of two patches to calculate intensity similarity. But because the contrast of an image is likely to be influenced by illumination, this method is not valid when the illumination condition changes between adjacent frames.

In order to reduce the influences brought by illumination, contrast of adjacent frames around illumination changes must be adjusted to make them have similar visual appearances. In our work, we first use histogram equalization as a way to adjust the contrast. Then, we discuss the scale to do the adjustment and how to measure the difference between a candidate patch and a reference patch.

Histogram equalization processing is defined as follows:

$$v_{out} = F_{hist}(v_{in}) = G^{-1}(T(v_{in})), \quad (3)$$

where $T(x) = \int_0^x p_x(w)dw$, $G(y) = \int_0^y p_y(t)dt$, and $p_x(x)$ is the probability density function of intensity level $x$, so as $p_y(y)$.

We have three scales to do the adjustment in NLM-based framework: patch scale, search window scale and image scale. Because illumination alters intensity of different areas of an image in different ways, adjusting contrast in image scale is not appropriate and it may bring error in local areas. Due to the number of pixels in a patch is too small, patch scale adjustment is sensitive to noise and inaccuracy. So we adjust contrast in search window scale considering that it has relatively adequate pixels and is adaptive to local illumination variance.

As the foreground and background of an image have different reflection under the same illumination condition, contrast adjustment may even bring serious errors when a search window contains both foreground and background object. To solve this problem, in our algorithm, instead of simply using the adjusted patch as the candidate patch to measure the intensity similarity, we do as follows:

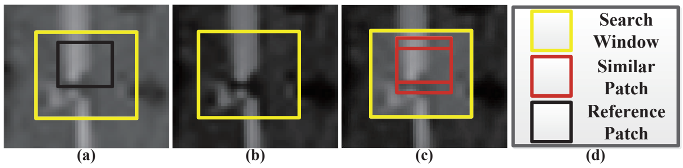


Fig. 1. (a) A reference patch in $k$-th frame. (b) Similar patches in $(k-1)$-th frame. (c) Similar patches in $(k-1)$-th frame by using our intensity measurement.

$$\begin{aligned} d_{int}(k,l,i,j) = min(&\|R_{k,l}Y_t - R_{i,j}Y_{t'}\|_2^2, \\ &\|R_{k,l}Y_t - R_{i,j}F_{histw}(Y_{t'})\|_2^2), \end{aligned} \quad (4)$$

where $F_{histw}$ is a filter which does histogram equalization in corresponding search window. This selection makes our measurement applicable to patches without illumination changes. So the intensity similarity between the candidate patch and the reference patch is defined as:

$$w_{int}(k,l,i,j) = \frac{1}{C_1(k,l)} \exp\left\{-\frac{d_{int}}{2\sigma_1^2}\right\}, \quad (5)$$

where $C_1(k,l)$ is defined as:

$$C_1(k,l) = \sum_{(i,j)\in N(k,l)} \exp\left\{-\frac{d_{int}}{2\sigma_1^2}\right\}.$$

The candidate patches whose $w_{int} > 0.1$ are defined as similar candidate patches of a reference patch. Fig.1 shows that when the illumination condition changes in the $k$-th frame, few similar candidate patches can be found in the $(k-1)$-th frames. But after adjusting contrast in search window scale, we can find more intensity similar patches between adjacent frames.

Finally, if we choose the adjusted patch as the candidate patch to measure intensity similarity, we do the adjustment in low-resolution image and pick the reference point in NLM-based framework from the adjusted low-resolution patch.

### B. Structure Similarity Measurement

Since histogram equalization is less effective when the contrast characteristics vary across the images, we may still not make the similar patch stand out in complex illumination condition even we adjust the contrast. In this situation, we need to consider other information to define similarity between patches besides intensity.

In complex illumination condition, the intensity, texture and many other attributes in adjacent frames may vary significantly, but the structure in local areas rarely changes. In our work, we extract local structure to obtain a structure descriptor. Then we calculate the structure similarity between different patches by using this descriptor.

Speed Up Robust Features (SURF) [3] is simplified to describe a local structure. We show the way to get the structure descriptor of a patch in Fig. 2. The important steps are shown as follows:

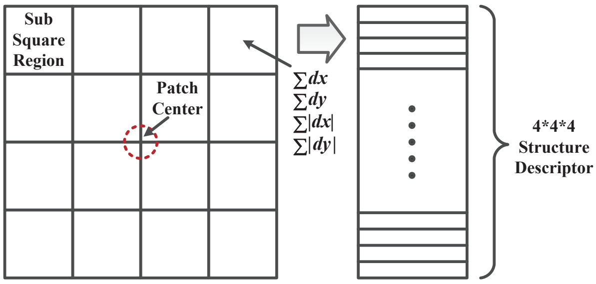


Fig. 2. The way to get the structure descriptor of a patch. Red circle is the center of a patch. Squares represent the sub-square-regions. The structure descriptor is the column vector on the right.

1) Build a square region around the center of a patch;
2) Split this region into 4*4 sub-square-region;
3) Calculate Haar wavelet responses in horizontal direction and vertical direction to get $d_x$ and $d_y$ in each sub-square-region;
4) Calculate $\sum d_x$, $\sum d_y$, $\sum |d_x|$, $\sum |d_y|$ in each sub-square-region;
5) Get a 4*4*4 descriptor and turn it into a unit vector to get illumination-invariance descriptor of a patch.

The Haar wavelet in different directions can represent important spatial information, so the descriptor contains structure information of a patch. By turning the original descriptor into a unit vector [2], we can make this descriptor invariant to illumination changes.

After we get structure descriptor of a patch, the structure similarity of two patches is defined as follows:

$$w_{str}(k,l,i,j) = \frac{1}{C_2(k,l)} \exp\left\{-\frac{\|D_{k,l} - D_{i,j}\|_2^2}{2\sigma_2^2}\right\}, \tag{6}$$

where $D_{k,l}$ represents structure descriptor of the patch which centers at $(k,l)$ and $C_2(k,l)$ is defined as:

$$C_2(k,l) = \sum_{(i,j)\in N(k,l)} \exp\left\{-\frac{\|D_{k,l} - D_{i,j}\|_2^2}{2\sigma_2^2}\right\}.$$

### C. Illumination-Invariance Similarity Measurement

We combine the intensity term and structure term together to get an illumination-invariance measurement to calculate similarity between candidate patch and reference patch. The illumination-invariance measurement is defined as follows:

$$w(k,l,i,j) = \frac{1}{C_3(k,l)} (w_{str}(k,l,i,j) + \beta \cdot w_{int}(k,l,i,j)), \tag{7}$$

where $\beta$ is used to balance structure term and intensity term and $C_3(k,l)$ is defined as:

$$C_3(k,l) = \sum_{(i,j)\in N(k,l)} (w_{str}(k,l,i,j) + \beta \cdot w_{int}(k,l,i,j)).$$

As we mentioned in Section III-B, contrast adjustment may be invalid in boundaries between background and foreground. So we should increase the weight of the structure term in the boundary areas. We observe that in the boundary areas, the $d_{int}$ of a patch is always small. So we decide the parameter $\beta$ according to the $d_{int}$ of a patch. $\theta$ is the threshold of $d_{int}$. The parameter $\beta$ of a patch is chosen as follows:

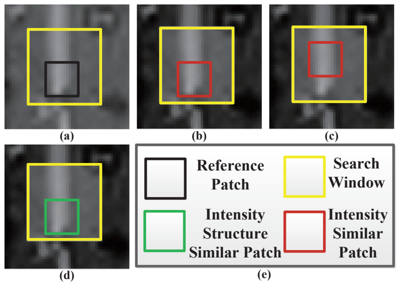


Fig. 3. (a) A reference patch in $k$-th frame. (b) (c) Similar patches in $(k-1)$-th frame by using our intensity measurement. (d) Similar patches by using illumination-invariance measurement.

$$\beta = \begin{cases} \lambda_1 & if \quad sum(d_{int}) < \theta, \\ \lambda_2 & if \quad sum(d_{int}) \geq \theta. \end{cases} \tag{8}$$

Fig. 3 shows that by making intensity information more accurate and adding extra structure information, our illumination-invariant measurement gets more reasonable results. The reference patch has a slight slope at the end of zip, as Fig. 3(a) shows. If we just use intensity measurement, we still consider the candidate path without the slope at the end of the zip as a similar patch (Fig. 3(b)). But after we use invariance measurement, we can eliminate this error, as Fig. 3(d) shows. Our illumination-invariance algorithm is summarized in Algorithm 1.

**Algorithm 1**: Illumination-Invariance Nonlocal Means SR.

**Input:** Input LR images $y_t(t = 1, \cdots, T)$ and input HR images $Y_t(t = 1, \cdots, T)$ generated by simple interpolation. Scale of SR reconstruction $s$. Size of patch in reference frame *block* and size of search window $winsz$. Number of reference frame $ref(1 \leq ref \leq T)$.
**Initialization:** Initialize $V, W$ with two zero matrixes whose size are equal to the HR reference frame.
**Contrast adjustment:**
For each $t \in [1, T]$, each $(i,j) \in t$-th HR images,
. Extract search window $W_{(i,j)}$, centering at $(i,j)$.
. Adjust contrast between search window $W_{(i,j)}$ and $W_{(k,l)}$.
**Weight processing:**
For each $(k,l) \in Y_{ref}$, $(i,j) \in N(k,l)$,
. Calculate $w_{int}(k,l,i,j)$ according to Eq.(5).
. Calculate $w_{str}(k,l,i,j)$ according to Eq.(6).
. Select weight parameter and calculate $w(k,l,i,j)$ according to Eq.(7).
. Update $V(k,l) += \Sigma_{t\in[1,\cdots,T]} \Sigma_{(i,j)\in N(k,l)} w(k,l,i,j) y_t(i,j)$.
. Update $W(k,l) += \Sigma_{t\in[1,\cdots,T]} \Sigma_{(i,j)\in N(k,l)} w(k,l,i,j)$.
**Reconstruction:**
For each $(k,l) \in Y_{ref}$, set $Res(k,l) = V(k,l)/W(k,l)$.
**Output:** HR image $Res(k,l)$.

## IV. Experimental Results

We conduct several experiments to verify the efficiency of our proposed method. In our experiment, the patch size is set to 7*7, the searching window is set to 21*21*3 (three adjacent frames), $\lambda_1 = 2$, $\lambda_2 = 1$, $\theta = 0.01$. We select optimal $\sigma_1 = 2.2$ and $\sigma_2 = 49$ in Eq.(7) to produce the best performance.

Once the parameters are selected, they keep fixed through our experiments.

We test our method on *Crew* sequence which contains local illumination changes caused by flashlight. First, for the purpose of demonstrating the efficiency of our intensity similarity measurement, we combine our intensity similarity measurement with NLM SR to get ISNLM and compare ISNLM to the original NLM. By comparing ISNLM to our proposed illumination-invariance NLM based method (IINLM), the validity of our illumination-invariance similarity measurement can be illustrated. The PSNR gains of IINLM and ISNLM over original NLM are illustrated in Fig.4. Since ISNLM adjusts contrast between adjacent frames, it produces better PSNR results at the frames with illumination changes (89th, 90th, 100th, 105th, 109th), and well maintains the performance of original NLM at frames in stable illumination condition. Because we add structure information to calculate patch similarity in IINLM, it significantly outperforms ISNLM in complex illumination conditions (89th, 90th, etc.). In stable illumination conditions, IINLM outperforms ISNLM and original NLM except in a few frames. The average PSNR gains of IINLM over ISNLM and original NLM in stable illumination conditions is 0.10dB and 0.18dB.

For the 89th, 90th, 100th, 105th, 109th frames, the SR results of original NLM, ISNLM, IINLM are given in Table I. For these frames in complex illumination condition, ISNLM improves original NLM by $0.1 - 0.2$dB. IINLM improves original NLM by $0.2 - 0.4$dB. In Fig.5, we compare the visual quality of these three methods for the 100th frame. Since we modify local contrast, compared to original NLM, ISNLM eliminates block artifacts. Because of consideration of structure information, compared to ISNLM, IINLM preserves more details in marginal regions, as can be seen in zoomed images.

TABLE I
PSNR (dB) OF SR RESULTS IN *Crew* SEQUENCE

| Frame | NLM | ISNLM | IINLM |
|---|---|---|---|
| 89th | 29.02 | 29.13 | **29.31** |
| 90th | 28.17 | 28.38 | **28.52** |
| 100th | 27.97 | 28.17 | **28.36** |
| 105th | 27.55 | 27.83 | **27.93** |
| 109th | 28.24 | 28.47 | **28.63** |

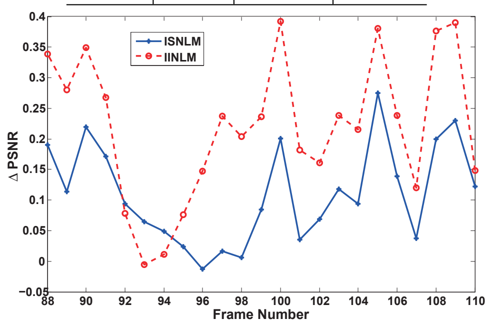


Fig. 4. The PSNR gains of the IINLM and ISNLM over original NLM on *Crew* sequence frame by frame.

Lastly, we run our method on a real video sequence, *Board*, captured by Canon PowerShot-A570 . We change the location

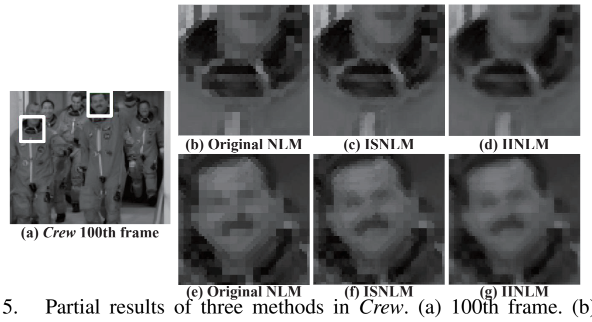


Fig. 5. Partial results of three methods in *Crew*. (a) 100th frame. (b) (e) Original NLM, (c) (f) ISNLM, (d) (g) IINLM.

and the intensity of the light source to stimulate a complex illumination condition. The 513th frame and zoomed details are shown in Fig.6. From all the experiments we have

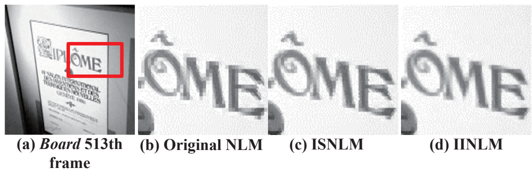


Fig. 6. Partial results of three methods in *Board*. (a) 513th frame. (b) Original NLM. (c) IINLM.

conducted, we can make a positive conclusion to our algorithm that illumination-invariant NLM SR is more reliable when processing the sequences in complex illumination condition and performs better than its basement, NLM SR, for obtaining more potential useful information from the candidate frames.

## V. CONCLUSIONS

NLM based SR algorithms are popular SR methods. But most of them are not adaptive to complex illumination conditions. In this paper, three possible improvements are proposed to handle this problem. We adjust the contrast between different search windows and select proper candidate patches. By incorporating structure information and combining structure and intensity information properly, our method produces robust results to illumination changes. Experiments show that our method is efficient in complex illumination conditions.

## REFERENCES

[1] M. Potter, M. Elad, H. Takeda, and P. Milanfar. Generalizing the Nonlocal-Means to Super-Resolution Reconstruction. *IEEE Transaction on Image Processing*, vol. 19, no. 1, pp. 36-51, Jan. 2009.

[2] D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.

[3] H. Bay, T. Tuytelaars, and L. V. Gool. SURF: Speeded Up Robust Features. *Europeon Conference on Computer Vision*, vol. 39, pp. 404-417, 2006.

[4] Y. Zhuo, J. Liu, J. Ren, and Z. Guo. Nonlocal Based Super Resolution with Rotation Invariance and Search Window Relocation. *International Conference on Acoustics, Speech, and Signal Processing*, Kyoto, Japan, Mar. 2012.

[5] M. Tanaka, Y. Yaguchi, and M. Okutomi. Robust and Accurate Estimation of Multiple Motions for Whole-Image Super-Resolution. *International Conference on Image Processing*, pp. 649-652, Oct. 2008.

[6] J. Ren, Y. Zhuo, J. Liu, and Z. Guo. Illumination-Invariant Non-Local Means based Video Denoising. *International Conference on Image Processing*, Sep. 2012.

[7] M.-H. Cheng, H.-Y. Chen, and J.-J. Leou. Video Super-Resolution Reconstruction using A Mobile Search Strategy and Adaptive Patch Size. *Signal Processing*, vol. 91, pp. 1284-1297, 2011.