Combining the Power of Internal and External Denoising

Inbar Mosseri* Maria Zontak* Michal Irani
Dept. of Computer Science and Applied Mathematics
The Weizmann Institute of Science, ISRAEL

Abstract

Image denoising methods can broadly be classified into two types: "Internal Denoising" (denoising an image patch using other noisy patches within the noisy image), and "External Denoising" (denoising a patch using external clean natural image patches). Any such method, whether Internal or External, is typically applied to all image patches. In this paper we show that different image patches inherently have different preferences for Internal or External denoising. Moreover, and surprisingly, the higher the noise in the image, the stronger the preference for Internal Denoising. We identify and explain the source of this behavior, and show that Internal/External preference of a patch is directly related to its individual Signal-to-Noise-Ratio ("PatchSNR"). Patches with high PatchSNR (e.g., patches on strong edges) benefit much from External Denoising, whereas patches with low PatchSNR (e.g., patches in noisy uniform regions) benefit much more from Internal Denoising. Combining the power of Internal or External denoising selectively for each patch based on its estimated PatchSNR leads to improvement in denoising performance.

1. Introduction

Image denoising is defined as the problem of recovering a natural image I from its noise-corrupted image I_N. This problem has a rich history, with considerable progress made in recent years. In particular, the idea of using recurrence of small image patches within a natural image for denoising was first introduced in [1], and later extended by [2, 7]). In these methods, each noisy image patch is denoised using other noisy patches within the noisy image. We refer to these as *"Internal Denoising"* methods. Other recent patch-based denoising methods employ external clean natural image patches (or a compact representation of them) to denoise each patch (e.g., [3, 9, 11]). We refer to these as *"External Denoising"* methods.

Internal and External denoising were previously compared by [10] in the context of the 'Non-Local Means' (NLM) settings [1]. They found Internal NLM to be superior to External NLM. However, their comparison assumed that the *same* method (Internal or External) was applied to *all* image patches. Applying the same method to all image patches is typically true for most denoising methods [1, 3, 4, 2, 7, 9, 11].

In this paper we show that some image patches inherently prefer Internal Denoising, whereas other patches inherently prefer External Denoising. *Combining the best of both should thus lead to better results.* We analyze and quantify the Internal vs. External preference of small noisy patches p_n (e.g., 7×7), and show that it is tightly related to the 'Signal-to-Noise-Ratio' within the patch, denoted by $PatchSNR(p_n)$. This is the ratio between the *empirical variance* of the original clean patch p, and the *empirical variance* of the noise within its noisy patch p_n.

We show that patches with low PatchSNR (e.g., in smooth image regions) tend to prefer Internal denoising, whereas patches with high PatchSNR (e.g., edges, texture) tend to prefer External denoising. The reason for this dichotomic behavior is the tradeoff between *noise-fitting vs. signal-fitting*. Patches with low PatchSNR are dominated by the noise. Unconstrained External search tends to overfit their noise (as opposed to constrained Internal denoising). On the other hand, patches with high PatchSNR (e.g., edges, texture) are dominated by the signal. They are less prone to overfitting the noise in an unconstrained External search, yet, they have much better signal-fit in the external clean database. These tradeoffs are analyzed and quantified.

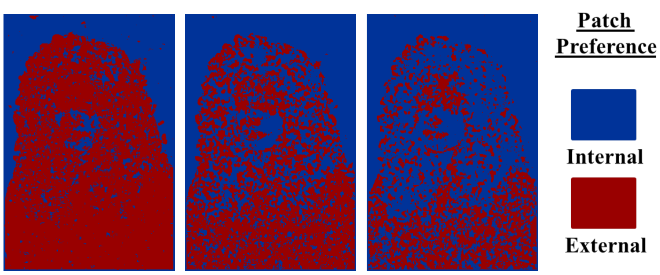
Finally, we exemplify the power of the PatchSNR, by combining pairs of different Internal/External denoising methods using a simple threshold on the patchSNR value (estimated per patch from the noisy image). Such a simple combination provides results better than the current state-of-the-art methods.

A recently published paper [5] also proposes to combine multiple denoising methods (up to 4). It is based on extensive machine-learning with *many parameters* (38 filter-response parameters computed per pixel), trained on hundreds of images (using Regression Tree Fields). In contrast, our goal is to understand *why* different patches prefer different methods. In particular, we show that much of the patch preference can be explained by a *single informative feature* - the PatchSNR. Moreover, we show that combining

*Equal contribution

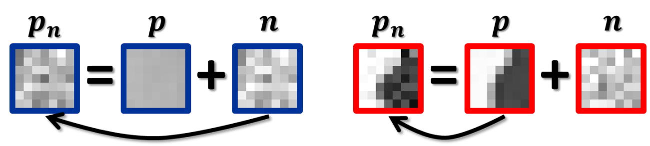
(a) Original (b) Noisy input ($\sigma = 35$) (c) Internal NLM (24.11dB) (d) External NLM (24.64dB) (e) Combining (c)&(d) using PatchSNR (25.60dB)

Figure 1: **Internal vs. External Denoising** *(NLM with 7x7 patches, $\sigma = 35$). Internal NLM is better for smooth patches; External NLM is better for detailed patches. Combining the two using a hard threshold on the PatchSNR value leads to a high quality image (e).*



(a) $\sigma = 15$ (b) $\sigma = 35$ (c) $\sigma = 55$

Figure 2: **Internal vs. External Patch Preference.** *Patch preference between Internal and External NLM for different noise levels: Red marks external preference; Blue marks internal preference. Note that the higher the noise in the image, the stronger the preference for internal denoising.*



(a) Patch dominated by noise (has low PatchSNR) (b) Patch dominated by signal (has high PatchSNR)

Figure 3: **Signal vs. Noise dominance within a patch.** *The noisy patches (a) and (b) are the Red and Blue patches marked in Fig. 1.b. Patch (a) prefers Internal NLM, whereas Patch (b) prefers External NLM.*

pairs of denoising methods, using a simple threshold on the PatchSNR parameter, provides results superior to current state-of-the-art methods (by up to 0.16dB). These results are equivalent to the results obtained in [5] when combining *pairs* of methods with *extensive multi-parameter training*. This conveys the importance of the single PatchSNR measure, and the cardinal role it plays in the denoising preference of patches.

The rest of the paper is organized as follows: Sec. 2 demonstrates the patch preferences for Internal vs. External denoising (visually and empirically). Sec. 3 analyzes and quantifies the tradeoff between Signal-fit and Noise-fit of a patch, and its relation to the PatchSNR. Sec. 4 suggests how to estimate the PatchSNR (per patch) from the noisy image. Sec. 5 provides experiments of simple combinations of a few different Internal and External denoising methods based on the estimated PatchSNR, with promising results.

2. Internal vs. External Patch Preference

In this section, we demonstrate the different preference of patches for Internal vs. External denoising. For simplicity, we exemplify and analyze these observations using the simple Non-Local-Means (NLM) setting [1]. However, the same observations hold also for more sophisticated state-of-the-art denoising methods, as will be shown in Sec. 5.

Figs. 1.c and 1.d show the denoising results of the Internal and External NLM (using 7x7 patches) on the noisy image of Fig. 1.b (Gaussian noise with $\sigma = 35$). Note that some parts of the image are better recovered by the Internal NLM (the smooth regions, e.g., the woman's face, the background), while other parts are better recovered by the External NLM (in particular, the textured regions, e.g., the woman's hair). Each noisy 7x7 patch was externally denoised by taking a weighted average of clean patches from an external database of 200 clean natural images (taken from the train set of the *BSDS300* [8]).

Fig. 2 further displays the Internal vs. External patch preference for three different noise levels (added to the image of Fig. 1.a). For each noisy patch we measured which of the two algorithms denoised it better (i.e., obtained a smaller RMSE) with respect to the ground truth image of

Fig. 1.a. Blue marks *internal preference*, Red marks *external preference*. Fig. 2 shows that the smooth patches prefer the Internal denoising, whereas patches with details (e.g. corners, edges, etc.) prefer the External denoising. Moreover, the higher the noise in the image, the stronger the preference for Internal denoising. The latter observation is quite surprising and counter-intuitive. One might expect that when the image patches are noisier, using clean patches (from an external database) would be preferable. Surprisingly, this is not the case.

To empirically validate that our observations hold in general for natural images, we repeated this experiment on 100 different natural images (taken from the test set of *BSDS300*) for 3 different noise levels ($\sigma = 15, 35, 55$). Fig. 4.a shows the average denoising error (RMSE) as a function of the patch variance (of the ground truth clean patch). It confirms that the preference for Internal NLM gets stronger as the noise level grows (the threshold between the Internal/External preference increases with the noise level σ). It further confirms that indeed patches with low variance (smooth or low-content patches) prefer Internal NLM, whereas patches with high variance (patches with details) prefer External NLM. *Combining the best of both should thus lead to better results.*

We further note that for patches which prefer Internal denoising (e.g., smooth patches), the noisy patch $p_n = p + n$ is dominated by the noise n, and not by the signal p (see Fig. 3.a). In other words, the patch p_n has low "Signal-to-Noise Ratio". On the other hand, for patches which prefer External denoising (e.g., edge, texture), the noisy patch p_n is dominated by the signal p (see Fig. 3.b). In other words, the patch p_n has high "Signal-to-Noise Ratio". We define the "Signal-to-Noise Ratio" (SNR) within a patch to be:

$$\text{PatchSNR}(p_n) \stackrel{\text{def}}{=\!=} \sqrt{\frac{var(p)}{var(n)}}$$

Fig. 4.b shows the RMSE per patch, but this time plotted as a function of the PatchSNR. As can be seen, there is a clear threshold between the Internal/External preference, which *does not depend on the global noise level σ*. This shows that the Internal/External denoising preference is tightly related to the PatchSNR. It further explains the phenomenon illustrated in Fig. 2: As the global noise level σ increases, the PatchSNR of each patch decreases. Therefore, fewer patches pass the PatchSNR threshold, thus more patches prefer Internal denoising over External. These observations will be analyzed and quantified in Sec. 3.

Fig. 1.e further shows a simple combination of Internal and External NLM (Figs. 1.c and 1.d) based on the PatchSNR: All patches whose PatchSNR was below $Threshold = 0.45$ (the threshold marked in Fig. 4.b), were denoised using Internal NLM, whereas all patches whose PatchSNR were above that threshold, were denoised using External NLM. The resulting image (Fig. 1.e) is of high quality. Not only is it significantly better than Figs. 1.c and 1.d, it is also significantly better than any of today's state-of-the-art denoising algorithms [2, 7, 11], achieving PSNR=25.60dB (compared to 24.50dB [2], 24.84dB [7], 24.85dB [11]). This is despite the fact that the resulting image was generated using the simple NLM denoising algorithm, with a simple one-parameter threshold. While not an algorithm (it relies on the *ground-truth* PatchSNR at each patch - information which is generally not available), this does convey the important role that the PatchSNR plays in determining Internal vs. External denoising preference. In the next section, we explain and analyze this tight relation between the 'Signal'-to-'Noise' Ratio within a patch, and its Internal vs. External denoising preference.

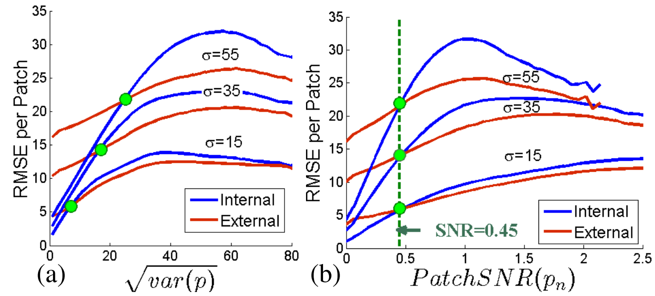


Figure 4: **Denoising Error as a Function of Patch-Variance and PatchSNR.** *Statistics performed on 100 natural images (noise levels $\sigma = 15, 35, 55$). Note that the threshold between Internal and External preference depends only on the PatchSNR, and is independent of the global noise level σ*

3. Noise-Fitting vs. Signal-Fitting

In this section we discuss the tradeoff between *signal fitting* and *noise fitting* for patches with different PatchSNRs. We show that in general, patches tend to have more and better representatives externally, in a large database of clean patches, than internally inside the same image (Sec. 3.1). However, the risk of overfitting the noise is also larger externally than internally (Sec. 3.2). We show that in the case of smooth patches, the overall tradeoff between signal-fit and noise-fit plays in favor of Internal denoising. In contrast, in the case of detailed patches, the overall tradeoff plays in favor of External denoising.

Since most existing denoising methods (Internal and External) are based on additive white Gaussian noise assumptions (e.g., [1, 4, 2, 9, 7, 11, 3]), our analysis will use the same underlying assumptions.

3.1. Fitting the Signal

In order to denoise a noisy patch $p_n = p + n$ *internally* (using other noisy patches), we need to average a sufficient number of similar patches. Assuming the averaged patches share a similar signal p, their independent added Gaussian

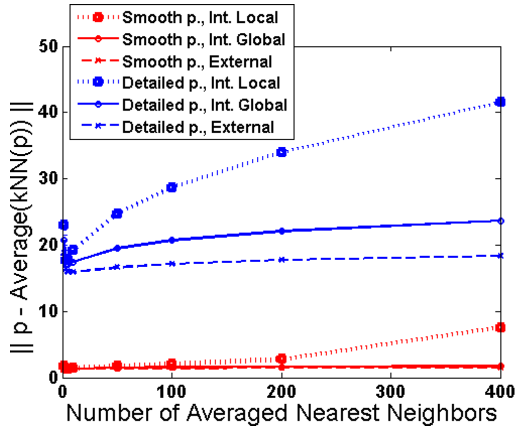


Figure 5: **Signal Fitting**: *RMSE between a clean patch p and the average of its k-NNs (computed for 7×7 patches).* ***Dotted curve (Internal-Local):*** *k-NNs selected from an internal 21×21 neighborhood around p;* ***Solid curve (Internal-Global):*** *k-NNs selected internally from the entire image;* ***Dashed curve (External):*** *k-NNs selected from an external database.* ***Red curves:*** *the 25% smoothest patches in the image;* ***Blue curves:*** *the 25% most detailed patches in the image. (Statistics over 100 images.)*

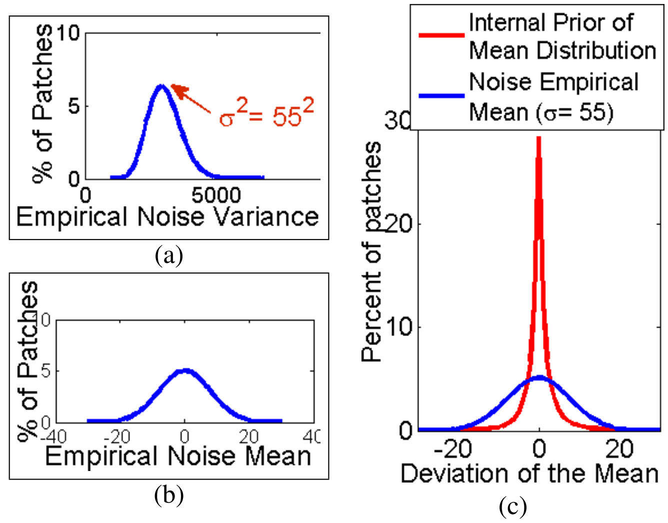


Figure 6: **Deviations of Empirical Mean and variance** *(shown for $\sigma = 55$): (a) Distribution of the empirical variance of 7×7 random noise patches. (b) Distribution of the empirical mean of 7×7 random noise patches. (c) Red curve - deviations of the mean of 7×7* clean *natural patches w.r.t. the central patch mean within a restricted 21×21 area. For comparison, the expected deviations of random noise mean are overlayed on top (marked in blue).*

noise n will be averaged out, allowing the clean signal p to emerge. Averaging k i.i.d. variables decreases the variance by factor k (*e.g.,* to reduce the noise standard-deviation σ by factor 10, we need to average 100 similar patches). When there are not enough good representatives of the signal p internally, we may prefer to go externally. As we will show next, smooth patches tend to have enough good representatives internally, whereas detailed patches do not.

Fig. 5 examines how well a clean "signal" p is represented internally vs. externally. The graph shows the RMSE between a clean patch p, and the average of its most similar patches, i.e., its k-Nearest Neighbors (k-NNs), computed for 7×7 patches. We examined three cases: (i) The k-NNs are selected from an internal 21×21 neighborhood surrounding p (dotted curve); (ii) The k-NNs are selected internally from the entire image (solid curve); (iii) The k-NNs are selected from an external database of 200 images (dashed curve). The red color refers to the 25% smoothest patches in the image (lowest var(p)), and the blue curve - to the 25% most detailed patches in the image (highest var(p)). These results were averaged over patches taken from 100 images.

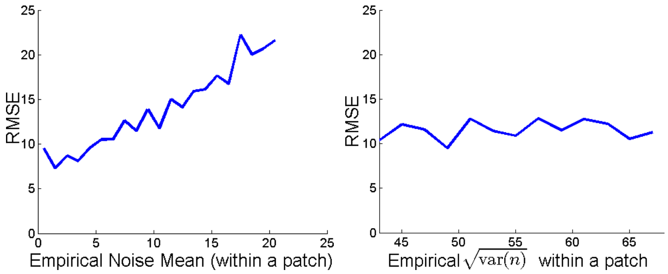
Note that the detailed patches p (blue curves) have much better representatives externally than internally. For example, averaging 100 internal NNs leads to a very high error. In fact, this error is much higher than averaging any number k of external NNs in the graph.

In contrast, smooth patches (red curves) have many good NNs both internally and externally. In both cases, averaging up to 100 NNs yields a reasonably good representative of the signal patch p, even in the case of a local internal search. Of course, once noise is added to p, the internal k-NNs (local or global) will also have noise, whereas the external ones will not, giving advantage to External denoising. However, averaging 100 internal noisy k-NNs will significantly reduce the noise (by a factor of 10). Moreover, as will be seen next, due to noise fitting problems, Internal denoising is preferable over External denoising in the case of smooth patches.

3.2. Overfitting the Noise

The noise is assumed to be additive random noise, sampled from Gaussian distribution with variance σ^2. Although the mean of the noise in the entire image is zero, in practice, the *empirical mean* (the *sample mean*) of the noise within an *individual* small patch (e.g., 7×7) is usually not zero (see Fig. 6.b). Similarly, the *empirical variance* (the *sample variance*) within an individual small patch is usually not σ^2 (see Fig. 6.a). We next analyze the noise-fitting in terms of these two factors: the empirical mean and the empirical variance of the noise within a small patch. In particular, we observe that most of the *inherent* denoising error reported in [6] for the case of *optimal* unconstrained External denoising of small patches, is due to *overfitting the non-zero mean of the noise* within the patch (and is invariant of the noise variance within the patch).

(a) **Overfitting the Noise-Mean**:
By assumption, the expected mean of the noise $\mathbb{E}(n)$ is zero.



(a) (b)

Figure 7: **Fitting of the Noise Mean:** *RMSE of the optimal external denoising of patches vs. empirical mean and variance of the noise within the patch (computed on the data of [6] for* 7×7 *patches,* $\sigma = 55$*). The denoising error grows linearly with the deviation from zero of the empirical noise-mean within the patch. In contrast, the denoising error is* independent *of the empirical noise-variance within the patch. (Average RMSE values are plotted)*

However, the empirical mean $\bar{n}$ of the noise component within a specific patch p_n is, in general, non-zero. Therefore, we can write: $p_n = p + n = \hat{p} + \hat{n}$, where $\hat{p} = p + \bar{n}$ and $\hat{n} = n - \bar{n}$. Noticeably, the empirical mean of $\hat{n}$ is now zero. Note that p and $\hat{p}$ are both natural image patches, which have the same details, but differ only in their means. Both decompositions of p_n are equally plausible. However, any typical denoising algorithm, which assumes zero mean noise, will tend to recover the more likely solution $\hat{p}$.

In other words, an *ideal* denoising algorithm (i.e., an algorithm that will recover $\hat{p}$ perfectly), will *inherently* have a residual error of $\hat{p} - p = \bar{n}$. Therefore, the expected error of an 'ideal' denoising algorithm, due to the difference between the empirical mean and the assumed zero mean, is

$$\text{RMSE}_{ideal} = \sqrt{\mathbb{E}_{\bar{n}}(\bar{n}^2)} = \sqrt{\int \bar{n}^2 P_r(\bar{n}) d\bar{n}} = \frac{\sigma}{\sqrt{d}} \ , \quad (1)$$

where d is the patch size, and $\bar{n} \sim N(0, \frac{\sigma^2}{d})$ is the noise mean within the patch. For example, for $\sigma = 55$ using 7×7 patches ($d = 49$), the minimal expected denoising error is RMSE$_{ideal}$=7.86 (can also be calculated empirically from Fig. 6.b). Denoising with smaller 5×5 patches will suffer even more from fitting the noise-mean (RMSE$_{ideal}$ = 11). These errors provide a *lower-bound* on the expected error in the case of *ideal* denoising given a patch size (which can be translated to an upper-bound on the expected PSNR performance).

We empirically verified that most of the error in *ideal* denoising of relatively small patches is due to overfitting the non-zero mean of the noise within a patch, using the data from [6] (kindly provided to us by the authors). In [6], upper-bounds for denoising algorithms are computed by denoising patches via *exhaustive* weighted average over a *huge* number of natural image patches (extracted from 20,000 clean natural images). This framework is close to an *ideal* denoising, because the chances to fit a signal p perfectly, are very high.

Fig. 7.a shows that the denoising error in the optimally-denoised data of [6] *grows linearly* with the deviation from zero of the noise-mean within the patch (shown for 7×7 patches, $\sigma = 55$). In contrast, the denoising error is *independent* of the empirical noise-variance within the patch (see Fig. 7.b). This confirms our observation that overfitting the non-zero noise-mean is inherent to unconstrained External denoising. For relatively small patches (e.g., 7×7), this is the major component of the residual denoising error in the optimally-denoised data of [6], and is the main source for their derived denoising bounds. For example, the error of relatively smooth patches ($\sqrt{var(p)} < 4$) from the data of [6] is RMSE=8.13, which is only slightly larger than the above estimated RMSE$_{ideal}$=7.86 *only* due to fitting the noise-mean.

The above derivations assumed a *uniform* prior on the mean of the clean patch p, which is true for *unconstrained External* denoising. In contrast - internally, when the denoising is restricted to 21×21 neighborhood surrounding a patch, there exists a *strong non-uniform prior* on the mean of the clean patch p. Patches in clean natural images tend to recur very densely in their immediate surrounding [10], hence their mean values tend to be very similar. This is especially true for uniform or low-content patches , which form a significant part of the image.

The red curve in Fig. 6.c displays the small *deviations* of patch mean values within 21×21 neighborhood in clean natural images (deviations are measured with respect to the mean value of the central patch). This distribution was empirically calculated for all relatively smooth 7×7 patches from 100 natural images of the BSD300 (with $\sqrt{var(p)} < 4$, which is roughly 1/3 of the patches). This distribution is highly localized around zero. Therefore, when restricting the denoising to averaging patches in a local neighborhood, the expected residual error due to fitting the non-zero noise mean, is much lower than in the unrestricted external search, especially for low-content patches. For example, even NLM denoising (which is far from being state-of-the-art) of these relatively smooth patches, obtains RMSE of 5.1 (in contrast to 8.13 obtained for these patches by the optimal External denoising of [6]). This is $\sim 4dB$ improvement of internal over external, for 1/3 of the patches!

(b) **Overfitting the Noise-Details:**
Although an ideal External denoising is independent of the empirical noise variance within a patch (Fig. 7.b), any practical algorithm, with a small limited external database, will not be. We next analyze effects of overfitting the "details" of the noise. To isolate the effects of detail-fitting, we first remove the mean of all the patches. For simplicity of notations, in the analysis below n, p, p_n, etc. will denote patches

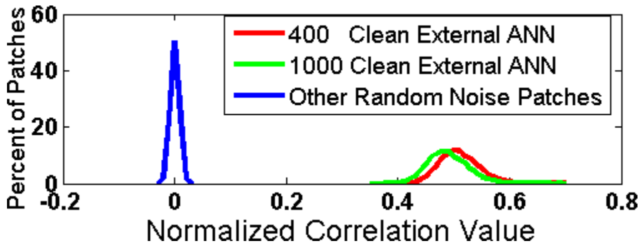


Figure 8: **Random-Noise Patches have High Normalized-Correlation with Natural Image Patches.** *Red and Green curves are distributions, showing high normalized-correlation values of random noise patches n with their 400 NNs among an external database of natural image patches (red curve), and 1000 NNs (green curve). In contrast, random-noise patches n have very low normalized-correlation with other random noise patches (Blue curve). Shown for 7×7 random noise patches n, for $\sigma = 55$.*

with zero mean.

Let us start from analyzing relatively smooth patches with low var(p). For such patches, after removing the mean, $p_n = p + n \simeq n$. These patches thus can be viewed as random noise patches. Let c denote the value of the Normalized-Correlation between a *random noise patch* n and its ''most similar" *natural patch* $NN(n)$. One would expect c to be low (closer to 0 than to 1). However, as shown in Fig. 8 (red and green curves), the value c is surprisingly high (around 0.5). Moreover, high normalized correlation values are obtained not only for the first Nearest Neighbor of n, but also for its other Nearest Neighbors (up to 1000 approximate nearest neighbors - ANN). For all our experiments, we used an external database of clean natural image patches from 200 images (from BSD300).

We experimented with several noise levels σ, and this correlation remains high (almost identical), regardless of σ. Finally, this correlation is even higher for smaller 5×5 patches (typically in the range of [0.5 0.8], with an average value of 0.65).

For every random noise patch we further calculated its ''denoised" version, by averaging its k-ANN in the external DB, using $k = 400$ and $k = 1000$. One would hope that such denoised patches would have very low variance (close to zero), due to averaging out the noise details. But this is not the case. Since all nearest neighbors are correlated to n, their average is also strongly correlated to n, leaving a non-negligable noise residues: $\|\text{Avg}(400NNs))\| = 16.46$; $\|\text{Avg}(1000NNs)\| = 15.49$ (where, $\|n\| = \sigma = 55$).

In contrast - internally, the denoising seeks other similar patches with independent random noise. When the Internal denoising is restricted to a local neighborhood of 20×20, denoising of smooth patches is equivalent to averaging 400 random noise patches (assuming same mean of the signal within the local neighborhood). The blue curve in Fig. 8 shows the distribution of normalized correlation values of 7×7 random noise patches ($\sigma = 55$), with other random noise patches. As expected, the average normalized correlation values are almost zero. The average of such 400 random noise patches yields a patch with zero empirical variance.

To summarize, relatively smooth patches are prone to overfitting of the noise-details, when unconstrained external search is performed. For such patches we would prefer an internal search in a restricted neighborhood, where we have no risk for noise fitting, but still have enough good signal representative (see Sec. 3.1).

We next analyze the case for more detailed patches. Let $q = NN(p_n)$ denote the nearest neighbor of $p_n = p + n$ in an external clean database. In general, the search for q will be guided both by the signal component p and by the noise component n. The question is, which of these two components will be more dominant in the search for $q = NN(p_n)$. As before, $NN(n)$ denotes the natural image patch in the external database, which is most similar to the noise component n of the patch p_n. Recall that c is the Normalized-Correlation value between n and $NN(n)$. It is easy to show that $\|NN(n)\| = c\|n\|$ (see proof in footnote[1]).

We next show that when the noisy patch p_n has low PatchSNR, then $NN(n)$ forms a *better 'nearest neighbor'* to p_n than the clean patch p. Let $dist_p$ denote the distance between the noisy patch p_n and its clean version p:

$$dist_p = \|p_n - p\|^2 = \|n\|^2.$$

Let $dist_n$ denote the distance between p_n and $NN(n)$:

$$\begin{aligned} dist_n &= \|p_n - NN(n)\|^2 \\ &= \|p\|^2 + \|n\|^2 + \|NN(n)\|^2 + \\ &\quad + 2\langle p, n\rangle - 2\langle p, NN(n)\rangle - 2\langle n, NN(n)\rangle \\ &\approx \|p\|^2 + (1 - c^2)\|n\|^2. \end{aligned}$$

The last step assumed that the signal p and the noise n are uncorrelated, hence their inner product (after having removed their means) is close to zero, hence negligible. The same holds for the inner product of p and $NN(n)$. However, the third inner product is far from being negligible, and is equal to: $\langle n, NN(n)\rangle = c \cdot \|n\| \cdot \|NN(n)\| = c^2\|n\|^2$. This term is quite large, as the normalized correlation value $c \approx 0.5$ (see Fig. 8 – red and green curves). Therefore:

$$dist_p > dist_n \iff \|p\| < c\|n\| \iff \frac{\|p\|}{\|n\|} < c.$$

Since all patch means have been removed: $\frac{\|p\|}{\|n\|} = \sqrt{\frac{\text{var}(p)}{\text{var}(n)}} = \text{PatchSNR}(p_n)$, defined in Sec. 2. Therefore:

$$dist_p > dist_n \iff \text{PatchSNR}(p_n) < c.$$

In other words, when PatchSNR(p_n)<c, an unconstrained external search for q=NN(p_n) will tend more

---

[1] *Proof*: NN(n) $= arg\min_r \|r - n\|^2$, s.t. c=NormCorr(n,NN(n)) $\Rightarrow$ $\|r - n\|^2 = \|r\|^2 + \|n\|^2 - 2\langle n, r\rangle = \|r\|^2 + \|n\|^2 - 2c\|r\|\|n\|$. Differentiating this expression w.r.t. $\|r\|$ and equating to 0, leads to: $\|NN(n)\| = \|r_{min}\| = c\|n\|$.

toward the noise NN(n) and lead to overfitting the noise details. Using a large number of Nearest Neighbors (e.g., 400, 1000) will not eliminate this problem (see Fig. 8 – red and green curves). However, when PatchSNR(p_n)>c, the search for q will tend more toward the signal p. In particular, as the SNR within the patch gets larger (e.g., for highly detailed patches, like edges, texture, with PatchSNR(p_n)≫c), the risk of fitting the noise details in an unconstrained external search is low. Yet, their 'signal-fitting' is much better externally than internally (Sec. 3.1). Thus, *External denoising* is preferable for detailed patches.

To summarize: Smooth patches, which are dominated by the noise, should prefer Internal denoising (where they have low risk of overfitting the noise-details or noise-mean, and have good enough signal-fit). In contrast, detailed patches, which are dominated by the signal, should prefer External denoising (where they have much better signal-fit, yet, do not risk fitting the noise-details). External denoising will inevitably suffer from fitting the noise-mean, *regardless of the PatchSNR*. However, for *detailed patches*, this error is substantially smaller than the signal-error introduced by their bad signal-fit in Internal denoising (compare the average signal-error in Fig. 5 (dotted blue line) to the average-error in Fig. 7.b). Thus, External vs. Internal denoising preference of a patch, should be determined by the signal vs. noise dominance within the patch, captured by its PatchSNR. This is also confirmed by the clear PatchSNR threshold in Fig. 4.b.

4. Estimating the Patch-SNR

So far, the PatchSNR was known to us – it was computed using the ground-truth clean image, I, and the noisy image, I_N. In practice, we only have the noisy image. In this section we propose how to estimate the PatchSNR from the noisy image. Let $p_n = p + n$ be a noisy patch. Assuming p and n are uncorrelated: $\text{var}(p_n) = \text{var}(p) + \text{var}(n)$. Thus:

$$\text{PatchSNR}(p_n) \quad = \quad \sqrt{\frac{\text{var}(p)}{\text{var}(n)}} = \sqrt{\frac{\text{var}(p_n)}{\text{var}(n)} - 1} \qquad (2)$$

While $\text{var}(p_n)$ is known (this is the empirical variance of the noisy patch), $\text{var}(n)$ is unknown and patch-dependent. The empirical variance of the noise in an individual noisy patch often deviates from the expected σ^2 (see Fig. 6.a).

We approximate $\text{var}(n)$ using one of the existing denoising algorithms, as follows: Let $\hat{p}$ be a denoised version of p_n. Then $\text{var}(n) \cong \text{var}(p_n - \hat{p})$. Note that the denoising here is not used for obtaining the final denoised patch $\hat{p}$, but only for estimating its empirical noise variance $\text{var}(n)$. In practice, we tried estimating $\text{var}(n)$ using two types of denoising algorithms: NLM [1] and BM3D [2]. We found the latter to provide better PatchSNR approximations (especially in the critical range where PatchSNR<1).

Fig. 9.a shows the result of using the estimated PatchSNR for combining Internal and External NLM. Patches with *Estimated_PatchSNR*≤0.45 were taken from Internal NLM (Fig. 1.c). Patches with *Estimated_PatchSNR*>0.45 were taken from External NLM (Fig. 1.d). The threshold 0.45 was determined by the dashed green line in Fig. 4.b. The resulting image (PSNR=25.49dB) is significantly better than either Internal or External denoising (PSNR=24.11dB and 24.64dB), and only slightly inferior to the ground-truth PatchSNR result of Fig. 1.a. For this image, the result is significantly better than other state-of-the-art methods (Fig. 9.b-d), despite using only simple NLM with a simple hard threshold.

5. Experiments

We experimented with a simple combination of three different Internal and External denoising methods: (i) Internal and External NLM, (ii) Internal BM3D [2] and External NLM, (iii) Internal BM3D and External EPLL [11].

We ran BM3D and EPLL using the implementations provided by the authors. For the Internal NLM we used 7x7 patches, and a 21x21 neighborhood (for all noise levels). For the External NLM, we used 50 external NNs per patch.

We combined the different methods using a simple threshold on the estimated PatchSNR (all those below a threshold were internally denoised, and all those above a threshold were externally denoised). We experimented both with a *hard threshold* and with a *soft threshold* with a small margin around it (inside which the internal and external denoised patches are linearly combined). We found that a soft threshold provides slightly better results. We evaluated our algorithm on 100 natural images taken from the test set of the *BSDS300*, contaminated with white Gaussian noise (with $\sigma = 25, 35, 45, 55$).

Results are reported in Table 1, and are also compared against three state-of-the-art methods: BM3D [2], LSSC [7], and EPLL [11]. The choice of good threshold and margin is dependent on the internal and external methods being combined (since different methods use different patch sizes, different sizes of local search regions, and different number of external representatives). For each combination of internal and external method we tuned these two parameters on a few images, and then applied them to a hundred different images (same threshold and margin to all noise levels).

Note that the combination of the Internal and External NLM is significantly better than either method alone (by at least 0.7dB for all noise levels). The combination of BM3D and External NLM improves over both algorithms (0.17-0.23dB over BM3D; 1.21-2.94dB over External NLM). Moreover, this combination of BM3D and External NLM provides results better than all the state-of-the-art methods (by 0.04-0.16dB over EPLL [11]).

External NLM is computationally intensive, since it requires searching for 50 NNs for each noisy patch in a large

σ	Combined Int. & Ext.NLM	Combined BM3D & Ext.NLM	Combined BM3D & EPLL	Ext. NLM	Int. NLM	BM3D [2]	LSSC [7]	EPLL [11]
55	24.86	**25.29**	**25.29**	22.35	24.18	25.11	25.10	25.13
45	25.73	**26.05**	**26.06**	23.70	25.01	25.82	25.90	25.93
35	26.84	**27.06**	**27.07**	25.34	26.11	26.89	26.98	26.98
25	28.36	28.51	**28.54**	27.43	27.66	28.35	28.46	28.47

Table 1: **Combining Internal & External Denoising Methods.** *Average PSNR values on 100 images (σ= 25, 35, 45, 55).*

(a) Combined Int&Ext NLM with *estimated* PatchSNR PSNR = 25.49dB

(b) LSSC [7] PSNR = 24.84dB

(c) EPLL [11] PSNR = 24.85dB

(d) BM3D [2] PSNR = 24.50dB

Figure 9: **Combined Internal & External NLM** *($\sigma = 35$). (a) Combining Internal and External NLM using the estimated PatchSNR (Threshold=0.45). (b)-(d) Current state-of-the-art methods denoising results.*

(a) Noisy Input ($\sigma = 55$)

(b) Combined BM3D & EPLL with *estimated* PatchSNR PSNR=27.40dB

(c) BM3D [2] PSNR=27.22dB

(d) EPLL [11] PSNR=27.18dB

Figure 10: **Combined BM3D (Internal Denoising) and EPLL (External Denoising).** *Smooth regions with low PatchSNR (e.g., the runway) are corrupted by EPLL, due to the noise-fitting problem of External Denoising, and are better recovered by BM3D. On the other hand, detailed patches with high PatchSNR (e.g., the markings on the wings) are better recovered by the external EPLL. Combining the two methods obtains the best of both.*

external database. The state-of-the-art method EPLL [11] can also be regarded as an *external* method, and is computationally much more efficient. Results of the combined Internal BM3D & External EPLL are also reported in Table 1. They are better than either BM3D and EPLL alone, and are also slightly better than the combined Internal BM3D & External NLM.

Fig. 10 shows an example of combined BM3D and EPLL. Note that smooth regions with low PatchSNR (e.g., the runway) are corrupted by the EPLL, due to the noise-fitting problem of External denoising. In contrast, BM3D (an Internal denoising method) does not suffer from the noise fitting problem and recovers the smooth regions well. On the other hand, detailed patches with high PatchSNR (e.g., the markings on the wings) are better recovered by the External EPLL than by the Internal BM3D. Combining the two methods obtains the best of both.

Note that the results reported above do not require any learning. It only requires setting 2 parameters: the threshold and the margin. These parameters can be either set manually, or tuned automatically on few images. Yet, our results are comparable to the extensive learning-based method of [5] when combining a pair of methods. Their method uses 38 responses per pixel extracted from lots of different filters, trained on hundreds of images. This implies the strength of the single PatchSNR feature.

Our simple combination algorithm (based on a single threshold) improves over BM3D by 0.17-0.23dB, even for high noise level (e.g., $\sigma = 55$). This is despite the fact that our additional information comes from denoising with small 7x7 patches and non-optimal external denoising. This delta improvement is due to our observations that Internal denoising is more suited for patches with low PatchSNR, whereas External denoising is more suited for patches with high PatchSNR. We believe that more sophisticated use of the PatchSNR can lead to a significant leap both in theoretical denoising bounds and in denoising results.

Acknowledgement: *The authors would like to thank Shahar Kovalsky and Alon Faktor for their helpful comments and discussions. This work was funded in part by the Israel Science Foundation, the Israel Ministry of Science, and the Citigroup Foundation.*

References

[1] A. Buadess, B. Coll, and M. J, M. A non-local algorithm for image denoising. In *CVPR*, 2005.

[2] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3d transform-domain collaborative filtering. *IEEE T-IP*, 16(8), 2007.

[3] M. Elad and M. Aharon. Image denoising via learned dictionaries and sparse representation. In *CVPR*, 2006.

[4] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE T-IP*, 54(12), 2006.

[5] J. Jancsary, S. Nowozin, and C. Rother. Loss-specific training of non-parametric image restoration models: A new state of the art. In *ECCV*, 2012.

[6] A. Levin and B. Nadler. Natural image denoising: Optimality and inherent bounds. In *CVPR*, 2011.

[7] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. In *ICCV*, 2009.

[8] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *ICCV*, 2001.

[9] S. Roth and M. Black. Fields of experts. *International Journal of Computer Vision*, 82(2), 2009.

[10] M. Zontak and M. Irani. The internal statistics of a single natural image. In *CVPR*, 2011.

[11] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. In *ICCV*, 2011.