Open access • Proceedings Article • DOI:10.1109/ICIP.2013.6738136

Dynamic super resolution of depth sequences with non-rigid motions — **Source link** ↗

Kassem Al Ismaeil, Djamila Aouada, Bruno Mirbach, Bjorn Ottersten

Institutions: University of Luxembourg

Related papers:

- A simultaneous method for 3D video super-resolution and high-quality depth estimation

- Sparse depth super resolution

- Image Registration for Super-Resolution

- Simultaneous Super-Resolution of Depth and Images Using a Single Camera

- Fast super-resolution on moving objects in video sequences

Share this paper: 🅕 🐦 in ✉

View more about this paper here: https://typeset.io/papers/dynamic-super-resolution-of-depth-sequences-with-non-rigid-4qxlsksz5q

DYNAMIC SUPER RESOLUTION OF DEPTH SEQUENCES WITH NON-RIGID MOTIONS

Kassem Al Ismaeil⋆   *Djamila Aouada⋆*   *Bruno Mirbach†*   *Björn Ottersten⋆*

⋆SnT - University of Luxembourg
{kassem.alismaeil, djamila.aouada, bjorn.ottersten}@uni.lu

†Advanced Engineering - IEE S.A.
bruno.mirbach@iee.lu

ABSTRACT

We enhance the resolution of depth videos acquired with low resolution time-of-flight cameras. To that end, we propose a new dedicated dynamic super-resolution that is capable to accurately super-resolve a depth sequence containing one or multiple moving objects without strong constraints on their shape or motion, thus clearly outperforming any existing super-resolution techniques that perform poorly on depth data and are either restricted to global motions or not precise because of an implicit estimation of motion. The proposed approach is based on a new data model that leads to a robust registration of all depth frames after a dense upsampling. The textureless nature of depth images allows to robustly handle sequences with multiple moving objects as confirmed by our experiments.

Index Terms— Depth sequence, dynamic super-resolution, motion estimation, upsampling, ToF data, moving object.

1. INTRODUCTION

Super resolution (SR) is the process of recovering a high resolution (HR) image from a set of captured low resolution (LR) frames. SR has originally been defined for static scenes, i.e., scenes where the motion between the observed images is global as opposed to dynamic scenes containing a moving object. The past two decades have witnessed tremendous work on SR for static scenes. As presented in [3], these algorithms, commonly referred to as classical SR, are numerically limited to small global motions even for an increased number of LR frames. Moreover, they cannot handle scenes with moving objects and consider the corresponding frames as outliers. As a solution to these major limitations, example-based SR algorithms have been proposed [4], as well as their combinations with classical multi-frame SR [5]. However, such algorithms depend on a heavy training phase and the quality of the super-resolved image is dependent on the suitability of the training data.

Relatively little attention has been given to the SR of dynamic scenes. Farsiu et al. have proposed in [1] a dynamic shift and add model (dynamic S&A) as a mere extension of the original S&A that was first defined for the static case in [2], hence suffering from the same restrictions. Other methods [6, 7, 8, 9] have been proposed to tackle the problem of dynamic SR by segmenting the moving object first before super resolving it. Such methods do not handle pixels on the boundary of the object causing major artifacts. In 2010, van Eekeren et al. [9] proposed an algorithm to solve the problem of boundary pixels; however, this algorithm is computationally heavy and based upon strong assumptions. In 2009, dynamic SR models were proposed with an implicit motion estimation, e.g., steering kernels for SR (SKSR) [21]. While the idea is theoretically attractive, it is very impractical as it relies on heavy computations and on many empirical parameters. Moreover, these methods are dedicated for 2D intensity sequences, strongly failing when it comes to depth data because of its abrupt value changes around edges and textureless nature. Such data, usually captured with a time-of-flight (ToF) camera, requires a resolution enhancement. Fusion based methods have been proposed as a solution for dynamic depth scenes [10, 11, 12, 13, 14] where a HR 2D camera is coupled with a depth LR camera. These methods often suffer from texture copying problems and require a perfect alignment and synchronization of 2D and depth sequences.

In [16], we have proposed to release the limitations on scale and motion of the static S&A algorithm [1, 2] in the case of depth data with global rigid motions. In this work, we target the SR of dynamic depth scenes containing one or multiple moving objects without prior assumptions on their shape or motion, and without engaging in an additional learning stage. To the best of our knowledge, we are the first to explore the multi-frame SR framework for dynamic depth scenes. The proposed algorithm takes advantage of the textureless nature of depth data, leading to a robust median estimation without fusing with 2D data; hence, avoiding blurring and texture copying artifacts. This algorithm is based on a new data model that starts by densely upsampling the LR measurements for an accurate registration using a new *cumulative motion compensation*.

The organization of the paper starts by formulating the problem of dynamic SR in Section 2. We then provide our key concepts for a robust motion estimation in Section 3. In Section 4, we propose a new data model that leads to a robust dynamic depth SR algorithm. In Section 5, we experimentally compare its performance with state-of-art techniques using depth sequences. A conclusion is given in Section 6.

2. PROBLEM FORMULATION

The aim of dynamic SR algorithms is to estimate a sequence of N HR images $\{\mathbf{x}_t\}_{t=1}^N$ of size $(m \times n)$ from an observed LR sequence $\{\mathbf{y}_t\}_{t=1}^N$, where each LR image $\mathbf{y}_t$ is of size $(m' \times n')$ pixels, with $n = r \cdot n'$ and $m = r \cdot m'$, such that r is the SR factor. Every image $\mathbf{y}_t$ may be viewed as an LR noisy and deformed realization of $\mathbf{x}_{t'}$ at the acquisition time t, with $t' \leq t$. Rearranging all images $\mathbf{x}_t$ and $\mathbf{y}_t$, $t = 1, \cdots, N$, in lexicographic order, i.e., column vectors of lengths mn and $m'n'$, respectively, we consider the following data model:

$$\mathbf{y}_t = \mathbf{DHL}_t^{t'}\mathbf{x}_{t'} + \mathbf{n}_t,\ t' \leq t \text{ and } t, t' \in [1, N] \subset \mathbb{N}^*,\quad (1)$$

where $\mathbf{D}$ is a matrix of dimension $(m'n' \times mn)$ that represents the downsampling operator, and which we assume to be known and constant over time. The system blur is represented by the time and space invariant matrix $\mathbf{H}$. The vector $\mathbf{n}_t$ is an additive Laplacian noise at time t as justified in [2]. The matrices $\mathbf{L}_t^{t'}$ are $(mn \times mn)$ matrices corresponding to the geometric motion between the considered HR image $\mathbf{x}_{t'}$ and the observed LR image $\mathbf{y}_t$ prior to its downsampling.

The dynamic SR problem is simplified by reconstructing one HR image at a time using the full observed sequence. From now on, we fix the reference time to t_0, and focus on the reconstruction of $\mathbf{x}_{t_0}$ from $\{\mathbf{y}_t\}_{t=t_0}^N$. The operation may be repeated for $t_0 = 1, \cdots, N$. Based on the data model in (1), and using an L_1 norm between the observations and the model, the Maximum Likelihood (ML) estimate of $\mathbf{x}_{t_0}$ is obtained as follows:

$$\hat{\mathbf{x}}_{t_0} = \arg\min_{\mathbf{x}_{t_0}} \sum_{t=t_0}^N \|\mathbf{DHL}_t^{t_0}\mathbf{x}_{t_0} - \mathbf{y}_t\|_1. \quad (2)$$

Using the same approach as in [2, 18], we consider that $\mathbf{H}$ and $\mathbf{L}_t^{t_0}$ are block circulant matrices. Therefore:

$$\mathbf{HL}_t^{t_0} = \mathbf{L}_t^{t_0}\mathbf{H}. \quad (3)$$

The minimization in (2) can therefore be decomposed into two steps; estimation of a blurred HR image $\mathbf{z}_{t_0} = \mathbf{Hx}_{t_0}$, followed by a deblurring step. In what follows, we assume that $\mathbf{y}_t$ is simply the noisy and decimated version of $\mathbf{z}_t$ without any geometric warp. We may thus write $\mathbf{L}_t^t = \mathbf{I}$, $\forall t$, $\mathbf{I}$ being the identity matrix, hence, $\mathbf{L}_t^{t_0}\mathbf{z}_{t_0} = \mathbf{z}_t = \mathbf{Hx}_t$. This operation can be assimilated to registering $\mathbf{z}_{t_0}$ to $\mathbf{z}_t$. We draw attention to the fact that in the case of static multi-frame SR, instead of a sequence, a set of observed LR images is considered, i.e., there is no order between frames. Such order becomes crucial in dynamic SR because the estimation of motion, based on the optical flow paradigm, happens between consecutive frames only. An accurate dynamic SR estimation is consequently highly dependent on the accuracy of estimating the registration matrices $\mathbf{L}_t^{t-1}$, as well as $\mathbf{L}_t^{t_0}$. In the case of one moving object with a very small translational motion

through few frames, a subpixel motion estimation would be sufficient to guarantee a good HR image. This assumption is not valid anymore if the object moves fast or the scene has multiple objects moving with different motions. In this case, the SR process becomes more challenging and a robust registration method is required using a dense optical flow. Most SR algorithms are directly related to a registration based on a too coarse pixel correspondence as compared to the scale of details in the scene. It is therefore necessary to call upon a very accurate *subpixel correspondence*. In what follows, we argue that this accuracy is highly increased after an upsampling of the observed sequence as presented in Section 3. We accordingly propose a new data formulation for dynamic depth SR and give its corresponding algorithm in Section 4.

3. MOTION ESTIMATION AND REGISTRATION

It has been shown in [17] that higher image resolutions help increase the accuracy of motion estimation which justifies applying an upsampling framework to get higher scale images. Moreover, performing the registration process on upsampled images guarantees a better result with a higher accuracy than registering the LR images $\mathbf{y}_t$ followed by upsampling them. This is due to the fact that registration parameters are approximated by rounding the motion vectors with an expected error of $\pm\frac{1}{2}$ pixel. The effect of this error is related to the size of the registered images, whereas the upsampling process reduces this effect from $\pm\frac{1}{2m}$ in the LR case to $\pm\frac{1}{2rm}$. Hence, we propose to upsample the observed LR images even before registering them. Due to the specifications of depth data, classical interpolation based methods (e.g. bicubic) cannot be used and lead to jagged values and blurring effects especially for boundary pixels. Thus, we propose to densely upsample $\mathbf{y}_t$, $t = 1, ..., N$, up to the super-resolved image of size $(m \times n)$. We define the resulting image as:

$$\mathbf{y}_t \uparrow\ = \mathbf{U} \cdot \mathbf{y}_t, \quad (4)$$

where $\mathbf{U}$ is a dense upsampling matrix of size $(mn \times m'n')$, which we choose to be the transpose of $\mathbf{D}$, s.t., $\mathbf{UD} = \mathbf{A}$, where $\mathbf{A}$ is a block circulant matrix that defines a new blurring matrix $\mathbf{B} = \mathbf{AH}$. Therefore, we redefine $\mathbf{z}_t$ as:

$$\mathbf{z}_t = \mathbf{Bx}_t. \quad (5)$$

Since the optical flow approach works under the assumption of small motions, the frames which are further from the reference frame $\mathbf{y}_{t_0} \uparrow$ would introduce a higher registration error than the ones that are closer to $\mathbf{y}_{t_0} \uparrow$. They will thus be considered as outliers. The percentage of these outliers is related to two main factors; the speed of the moving objects and the length of the sequence N. For example, a long sequence with a fast moving object would most likely lead to more than 50% of outliers and the SR process fails even when using a robust estimator with a high breakdown value such as a median estimator. To tackle this problem, we herein propose to use a new

registration method based on a *cumulative motion compensation*.

Considering two consecutive upsampled frames $\mathbf{y}_{t-1} \uparrow$ and $\mathbf{y}_t \uparrow$, the optimal registration solution is:

$$\hat{\mathbf{M}}_{t-1}^t = \arg \min_{\mathbf{M}} \Psi \left(\mathbf{y}_{t-1} \uparrow, \mathbf{y}_t \uparrow, \mathbf{M} \right), \qquad (6)$$

where Ψ is a dense optical flow-related cost function and

$$\mathbf{y}_t \uparrow = \mathbf{M}_{t-1}^t \mathbf{y}_{t-1} \uparrow + \mathbf{v}_t. \qquad (7)$$

The vector $\mathbf{v}_t$ contains the innovation that we assume negligible in this framework. In addition, similarly to [21], for analytical convenience, we assume that all pixels in $\mathbf{y}_t \uparrow$ originate from pixels in $\mathbf{y}_{t-1} \uparrow$ in a one to one mapping. Therefore, each row in $\mathbf{M}_{t-1}^t$ contains 1 for each position corresponding to the address of the source pixel in $\mathbf{y}_{t-1} \uparrow$. This bijective property implies that the matrix $\hat{\mathbf{M}}_{t-1}^t$ is an invertible permutation, s.t., $[\hat{\mathbf{M}}_{t-1}^t]^{-1} = \hat{\mathbf{M}}_t^{t-1}$. Furthermore, its estimate leads to the following registration to $\mathbf{y}_{t-1}$:

$$\overline{\mathbf{y}}_t \uparrow = \hat{\mathbf{M}}_t^{t-1} \mathbf{y}_t \uparrow . \qquad (8)$$

We then need to define $\overline{\mathbf{y}}_t^{t_0} \uparrow$, the registered version of $\mathbf{y}_t \uparrow$ to the reference $\mathbf{y}_{t_0} \uparrow$. To that end, we use all the registered upsampled images $\overline{\mathbf{y}}_t \uparrow$, as defined in (8), for $t > t_0$. We propose, similarly to our work in [17], a *cumulative motion compensation* approach with an additional improvement where we further reduce the cumulated motion error by recomputing $\hat{\mathbf{M}}_{t-1}^t$ as follows:

$$\hat{\mathbf{M}}_{t-1}^t = \arg \min_{\mathbf{M}} \Psi \left(\mathbf{y}_{t-1} \uparrow, \overline{\mathbf{y}}_t \uparrow, \mathbf{M} \right), \qquad (9)$$

We prove by induction the following registration relationship for non-consecutive frames:

$$\overline{\mathbf{y}}_t^{t_0} \uparrow = \hat{\mathbf{M}}_t^{t_0} \mathbf{y}_t \uparrow = \underbrace{\hat{\mathbf{M}}_{t_0+1}^{t_0} \cdots \hat{\mathbf{M}}_t^{t-1}}_{(t-t_0) \text{ times}} \cdot \mathbf{y}_t \uparrow . \qquad (10)$$

Considering the bijection simplification, we further write:

$$\hat{\mathbf{M}}_t^{t_0} \approx \mathbf{L}_{t_0}^t = [\mathbf{L}_t^{t_0}]^{-1}. \qquad (11)$$

4. PROPOSED ALGORITHM

The subpixel accuracy in motion estimation induced by the combined upsampling and cumulative motions proposed in Section 3, make it feasible to handle a depth sequence with a moving object without using any prior information on its shape, rigidity, or motion. These advantages are extended to the much more complex case of multiple moving objects. Indeed, the textureless nature of depth images categorizes them as images containing gross information only, i.e., with no texture information, as per Mallikarajuna et al.'s composite image model [22]. This property, combined with the SR impulse

noise $\mathbf{n}_t$, suggests that a temporal median estimator is a robust equivalent to the ML formulation of (2).

We reformulate the data model in (1) to introduce the upsampling strategy of Section 3. Combining (1), (3), (4), (5), and (11), we find[1]:

$$\overline{\mathbf{y}}_t^{t_0} \uparrow = \mathbf{z}_{t_0} + \mathbf{w}_t, \ t_0 \le t \text{ and } t, t_0 \in [1, N] \subset \mathbb{N}^*, \quad (12)$$

where $\mathbf{w}_t = \hat{\mathbf{M}}_t^{t_0} \mathbf{U} \cdot \mathbf{n}_t$ is an additive Laplacian noise at t. The estimation in (2) becomes:

$$\hat{\mathbf{z}}_{t_0} = \arg \min_{\mathbf{z}_{t_0}} \sum_{t=t_0}^{N} \|\mathbf{z}_{t_0} - \overline{\mathbf{y}}_t^{t_0} \uparrow \|_1, \qquad (13)$$

which corresponds to the pixel-wise temporal median estimator, i.e., $\hat{\mathbf{z}}_{t_0} = \text{med}_t \{\overline{\mathbf{y}}_t^{t_0} \uparrow\}_{t=t_0}^{N}$. Then follows a simple image deblurring to recover $\hat{\mathbf{x}}_{t_0}$ from $\hat{\mathbf{z}}_{t_0}$. We hence propose a new dynamic SR algorithm corresponding to the presented new SR estimation, that we refer to as *Upsampling for Precise Super-Resolution* (UP-SR) as summarized below:

UP-SR algorithm

---

for t_0,
1.　　Choose the reference frame $\mathbf{y}_{t_0}$.
for t, *s.t.*, $t_0 < t < N$,
do
2.　　Compute $\mathbf{y}_t \uparrow$ using (4).
3.　　Estimate the registration matrices $\hat{\mathbf{M}}_t^{t_0}$ using (10).
4.　　Compute $\overline{\mathbf{y}}_t^{t_0} \uparrow$ using (10).
end do
end for
5.　　Find $\hat{\mathbf{z}}_{t_0}$ by applying a median estimator (13).
6.　　Deduce $\hat{\mathbf{x}}_{t_0}$ by deblurring.
end for

---

5. EXPERIMENTS

We test the performance of the proposed UP-SR algorithm on depth data acquired with a ToF camera. Using the SR estimation on such data is suitable as it suffers from a very low resolution. We start with a simple case of one moving object (*hand*) with a translational motion. We compare the performance of UP-SR for both cases, registered measured LR depth images $\overline{\mathbf{y}}_t^{t_0}$ and registered densely upsampled depth images $\overline{\mathbf{y}}_t^{t_0} \uparrow$, $t_0 < t < N$. Results show that in the latter case (Fig. 1(b)), the registration is more accurate and leads to sharper edges. Directly relying on LR images, however, leads to blurred edges (Fig. 1(a)), necessitating a special treatment or a segmentation step to reduce the artifacts caused by the boundary pixels. This experimentally confirms the benefit of our upsampling strategy. Next, we tested UP-SR on a real se-

---

[1]A full proof of (12) will be provided in another paper.
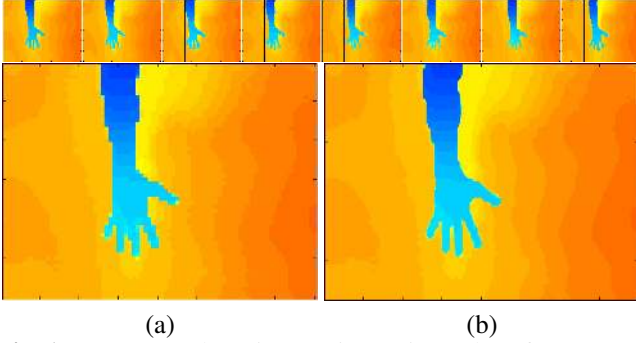
(a)   (b)

Fig. 1. UP-SR results using motion estimated (a) from an LR sequence; (b) from densely upsampled sequence ($r = 5$).

quence of LR depth images containing multiple moving objects. We mounted an LR ToF camera at a 2.5 meter hight looking down at the ground with two persons sitting on chairs sliding in two different directions. A sequence of 9 LR depth images, of size (56×61) pixels, were super-resolved with a factor $r = 5$ using UP-SR, 2D/depth fusion [13], SKSR [18] and dynamic S&A [1]. Visual results for one frame are given in Fig. 2(c), (d), (e) and (f), clearly showing that SKSR and dynamic S&A fail badly on depth data mainly on boundary pixels while 2D/depth fusion, although computationally efficient, often suffers from strong 2D texture copying on the final super-resolved depth frame. Fig. 2(f) shows the result of UP-SR where we obtained clear sharp edges in addition to an efficient removal of noisy pixel values. This is mostly due to the proposed subpixel motion estimation combined with an accurate registration leading to a successful temporal fusion of the sequence. Finally, in order to provide a quantitative evaluation, we generated an LR depth sequence by downsampling an available HR depth sequence with a factor $r = 4$, and further degrading it by additive white Gaussian noise (AWGN) with signal to noise ratio (SNR) of 15, 25, 35, and 45 dB. We quantitatively compare our proposed algorithm with SKRS and dynamic S&A. We tested these methods using the corresponding softwares provided in [19] and [20]. Since we have a known ground truth $\{\mathbf{x}_t\}_{t=1}^N$, we measure the quality of an estimated HR depth frame $\hat{\mathbf{x}}_{t'}$ using peak SNR (PSNR), which is defined as: PSNR $= 10 \log_{10} \frac{m \times n}{\|\mathbf{x}_{t'} - \hat{\mathbf{x}}_{t'}\|_2}$. Obtained results show the superiority of the UP-SR algorithm where it provides the best results among discussed state-of-the-art SR methods across all noise levels. As illustrated in Fig. 3, it is not surprising to see that even for a very high noise level (SNR = 15 dB) results are good. This is due to the key components of UP-SR, namely, its subpixel motion estimation and accurate multi-frame registration combined with a robust median filtering that matches the textureless property of depth data. Therefore, our algorithm results with good quality depth images without having to call upon an additional regularization step. The same algorithm applied on 2D images gives blurry results with lost details due to the fact that generally they do not fall under the model proposed in (12).
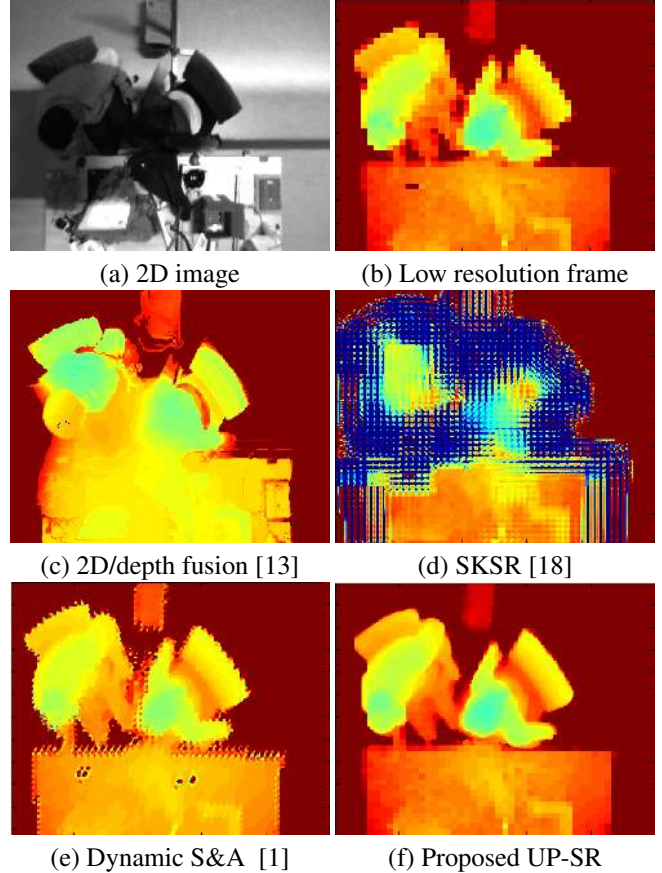

(a) 2D image   (b) Low resolution frame

(c) 2D/depth fusion [13]   (d) SKSR [18]

(e) Dynamic S&A [1]   (f) Proposed UP-SR

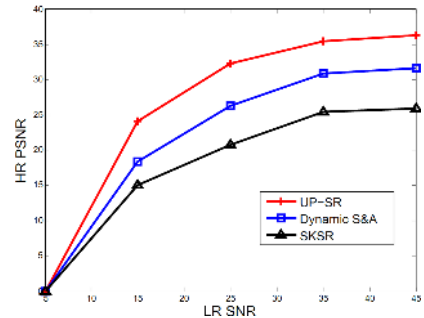**Fig. 2**. UP-SR example of a dynamic depth scene ($r = 5$).

Fig. 3. PSNR on the *moving hand* sequence with $r = 4$.

6. CONCLUSION

A new algorithm, UP-SR, has been presented to enhance the quality of LR depth videos for dynamic scenes containing one or multiple moving objects. This algorithm is based on the SR framework without strong constraints on objects' shape or motion. It takes advantage of the textureless nature of depth data to achieve robust SR estimation after densely upsampling LR frames. Experimental results with both synthetic and real ToF depth images showed that UP-SR, although conceptually simple, provides a more accurate motion estimation which leads to greatly outperforming existing methods.

7. REFERENCES

[1] Farsiu, S.; Robinson, M.D.; Elad, M.; Milanfar, P., "Advances and challenges in super-resolution", *International Journal of Imaging Systems and Technology*, vol. 14, pp. 47-57, 2004

[2] Farsiu, S.; Robinson, M.D.; Elad, M.; Milanfar, P., "Fast and robust multiframe super resolution," *Image Processing, IEEE Transactions on* , vol.13, no.10, pp.1327,1344, Oct. 2004

[3] Zhouchen Lin; Heung-Yeung Shum, "Fundamental limits of reconstruction-based superresolution algorithms under local translation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol.26, no.1, pp.83,97, Jan. 2004

[4] Aodha, Oisin Mac; Campbell, Neill D. F.; Nair, Arun; Brostow, Gabriel J., "Patch based synthesis for single depth image super-resolution," *Proceedings of the 12th European Conference on Computer Vision*, vol. Part III, ECCV'12, pp.71-84, ISBN: 978-3-642-33711-6

[5] Glasner, D.; Bagon, S.; Irani, M., "Super-resolution from a single image," *Computer Vision, 2009 IEEE 12th International Conference on*, vol., no., pp.349,356, Sept. 29 2009-Oct. 2 2009

[6] Hardie, R.C.; Tuinstra, T.R.; Bognar, J.; Barnard, K.J.; Armstrong, E., "High resolution image reconstruction from digital video with global and non-global scene motion," *Image Processing, 1997. Proceedings., International Conference on*, vol.1, no., pp.153,156 vol.1, 26-29 Oct 1997

[7] Farsiu, S.; Elad, M.; Milanfar, P., "Video-to-video dynamic super-resolution for grayscale and color sequences," *EURASIP Journal on Applied Signal Processing*, vol. 2006, pp.232-232, ID 61859

[8] van Eekeren, A.; Schutte, K.; Dijk, J.; de Lange, D.J.J.; van Vliet, L.J., "Super-Resolution on Moving Objects and Background," *Image Processing, 2006 IEEE International Conference* on , vol., no., pp.2709,2712, 8-11 Oct. 2006

[9] Van Eekeren, A. W M; Schutte, K.; van Vliet, L.J., "Multiframe Super-Resolution Reconstruction of Small Moving Objects," *Image Processing, IEEE Transactions on*, vol.19, no.11, pp.2901,2912, Nov. 2010

[10] J. Diebel and S. Thrun, "An application of Markov random fields to range sensing," *Proceedings of Conference on Neural Information Processing Systems (NIPS)*, pp. 291-298, 2005.

[11] Johannes Kopf and Michael F. Cohen and Dani Lischinski and Matt Uyttendaele, "Joint bilateral upsampling," *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2007)*, vol.26, no.3, 2007

[12] Qingxiong Yang; Ruigang Yang; Davis, J.; Nister, D., "Spatial-depth super resolution for range images," *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, vol., no., pp.1,8, 17-22 June 2007

[13] Garcia, F.; Aouada, D.; Mirbach, B.; Solignac, T.; Ottersten, B., "Real-time hybrid ToF multi-camera rig fusion system for depth map enhancement," *Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2011 IEEE Computer Society Conference on , vol., no., pp.1,8, 20-25 June 2011

[14] Xueqin Xiang; Guangxia Li; Jing Tong; Zhigeng Pan, "Fast and simple super resolution for range data," *Cyberworlds (CW), 2010 International Conference on* , vol., no., pp.319,324, 20-22 Oct. 2010

[15] Garcia, F.; Aouada, D.; Mirbach, B.; Ottersten, B., "Spatio-temporal ToF data enhancement by fusion," *Image Processing (ICIP), 2012 19th IEEE International Conference on* , vol., no., pp.981,984, Sept. 30 2012-Oct. 3 2012

[16] Al Ismaeil, K.; Aouada, D.; Mirbach, B.; Ottersten, B., "Depth super-resolution by enhanced shift & add, " Proceedings of the 15th International Conference on Computer Analysis of Images and Patterns, CAIP'13, York, UK, 27-29 Aug. 2013.

[17] L. Xu, J. Jia, S. B. Kang, "Improving sub-pixel correspondence through upsampling", *J. of Computer Vision and Image Understanding (CVIU)*, vol 116, Issue 2, February 2012, pp. 250-261, ISSN 1077-3142.

[18] Takeda, H.; Milanfar, P.; Protter, M.; Elad, M., "Super-resolution without explicit subpixel motion estimation," *Image Processing, IEEE Transactions on*, vol.18, no.9, pp.1958,1975, Sept. 2009

[19] http://users.soe.ucsc.edu/htakeda/SpaceTimeSKR.htm

[20] http://users.soe.ucsc.edu/milanfar/software/superresolution.html

[21] Elad, M.; Feuer, A., "Super-resolution reconstruction of continuous image sequences," *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on*, vol.3, no., pp.459,463 vol.3, 1999

[22] Mallikarjuna, H. S.; Chaparro, L.F., "Iterative composite filtering for image restoration," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol.14, no.6, pp.674,678, Jun 1992