(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 October 2010 (28.10.2010)

PCT

(10) International Publication Number
WO 2010/122502 A1

(51) **International Patent Classification**:
G06K 9/32 (2006.01)

(21) **International Application Number**:
PCT/IB2010/051738

(22) **International Filing Date**:
20 April 2010 (20.04.2010)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
61/170,884 20 April 2009 (20.04.2009) US
61/313,837 15 March 2010 (15.03.2010) US

(71) **Applicant** *(for all designated States except US)*: **YEDA RESEARCH AND DEVELOPMENT CO. LTD.** [IL/IL]; At The Weizmann Institute of Science, P.O. Box 95, 76100 Rehovot (IL).

(72) **Inventors; and**

(75) **Inventors/Applicants** *(for US only)*: **IRANI, Michal** [IL/IL]; 7 Neve Matz St., Weizmann Institute, 76100 Rehovot (IL). **GLASSNER, Daniel** [IL/IL]; c/o Yeda Research & Development Co. Ltd., At the Weizmann Institute of Science, POB 95, 76100 Rehovot (IL). **SHAHAR, Oded** [IL/IL]; 40 HaAvoda St., 63821 Tel-Aviv (IL). **BAGON, Shai** [IL/IL]; c/o Yeda Research & Development Co. Ltd., At the Weizmann Institute of Science, POB 95, 76100 Rehovot (IL).

(74) **Agent**: **SWIRSKY, Daniel J.**; AlphaPatent Associates Ltd., 55 Reuven St., 99544 Beit Shemesh (IL).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ,

[Continued on next page]

(54) **Title**: SUPER-RESOLUTION FROM A SINGLE SIGNAL

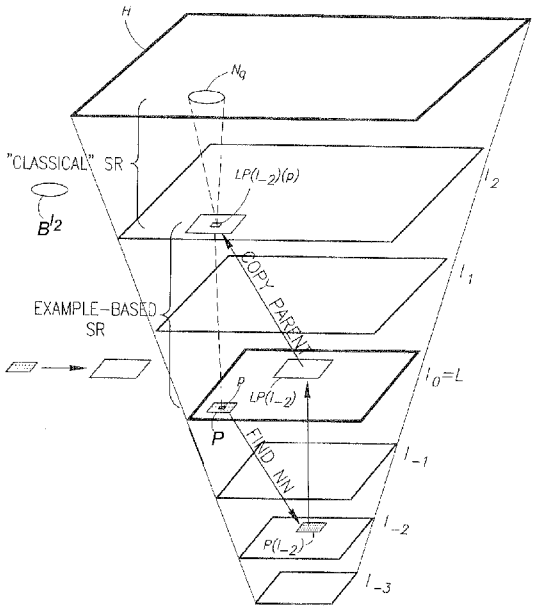


FIG.4B

(57) **Abstract**: A method implementable on a computing device includes exploiting data redundancy to combine high frequency information from at least two different scales of an input signal to generate a super resolution version of said input signal. An alternative method includes exploiting recurrence of data from an input signal in at least two different scales of at least one reference signal to extract and to combine high frequency information from a plurality of scales of said at least one reference signal to generate a super resolution version of said input signal. An alternative method includes generating a super resolution version of a single input video sequence in at least the temporal dimension by exploiting data recurrence within the input video sequence or with respect to an external database of example video sequences. A signal may be an image, a video sequence, an audio signal, etc.

WO 2010/122502 A1

TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

SUPER-RESOLUTION FROM A SINGLE SIGNAL

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims benefit from U.S. Provisional Patent Application 61/170,884, filed April 20, 2009, which is hereby incorporated in its entirety by reference.

FIELD OF THE INVENTION

[0002] The present invention relates to signal processing generally and to super-resolution of signals in particular.

BACKGROUND OF THE INVENTION

[0003] The goal of super-resolution methods is to recover a high resolution signal from one or more low resolution input signals. Methods for super-resolution of images can be broadly classified into two families of methods: the classical multi-image super-resolution and example-based super-resolution.

[0004] Reference is now made to Fig. 1, which generally illustrates the two types of methods.

[0005] In classical multi-image super-resolution such as described in Michal Irani and Shmuel Peleg, "Improving resolution by image registration", *CVGIP: Graphical Model and Image Processing,* Vol. 53,(3), 1991, pp 231-239, Capel, David, "Image Mosaicing and Super-Resolution", *Springer Verlag, 2004*, and Farsiu, S., Robinson, MD, Elad, M. and Milanfar, P., "Fast and robust multiframe super resolution", *IEEE Transactions on Image Processing,* Vol. 13, (10), 2004, pp 1327-1344, a set of low-resolution images L_i of the same scene are taken, at sub pixel misalignments, and the method attempts to generate a high-resolution image H from the multiple low-resolution images L_i.

[0006] Each pixel p of the low resolution image L_j is assumed to be generated from high resolution image H by a blurring function B_j and a sub-sampling process s_j, as given by Equation 1.

$$L_j = (H * B_j)\downarrow_{s_j} \qquad \text{Equation 1}$$

[0007] Thus, each low resolution pixel p in each low resolution image L_j induces one linear constraint on the unknown high-resolution intensity values within a local neighborhood

R around its corresponding high-resolution pixel q, where the size of the neighborhood is determined by the "support" of the blur kernel B_j:

[0008] $$L_j(p) = (H * B_j)(q) = \sum_{q_i \in Support(B_j)} H(q_i) B_j(q_i - q) \qquad \text{Equation 2}$$

[0009] where $H(q_i)$ are the unknown high-resolution intensity values and B_j is the blur kernel relating high resolution image H with low resolution image L_J. Thus, region $R_i(p)$, the region which affects pixel p in image L_j, is a function of p, the alignment between L_j and H at sub-pixel accuracy and the blur kernel B_j. A blurring function may be a point spread function and the amount of blur typically is a function of the extent of the sub-sampling, where the more the sub-sampling, the larger the blur.

[0010] Thus, each pixel $L_i(p)$ of low resolution image L_i imposes a constraint on local region $R_i(p)$. If enough low-resolution images are available (at subpixel shifts from each other), then the set of equations becomes determined and can be solved to recover the high-resolution image.

[0011] Such super-resolution schemes have been shown to provide reasonably stable results up to a factor of about 2, but are limited in the presence of noise and misregistration. These limitations have lead to the development of ``Example-Based Super-Resolution" introduced in articles by Freeman, W.T. and Pasztor, E.C. and Carmichael, O.T., "Learning low-level vision", *International Journal of Computer Vision,* Vol. 40 (1), 2000, Springer, pp 25-47, Freeman, William T., Jones, Thouis R. and Pasztor, Egon C.," Example-Based Super-Resolution", *IEEE Comput. Graph. Appl.,* Vol. 22, March 2002, (2), pp. 56--65, and Simon Baker and Takeo Kanade, "Limits on super-resolution and how to break them", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 2002, Vol. 24, pp. 1167—1183 and extended later by others.

[0012] In example-based super-resolution, pairs of corresponding images L and H (generally with a relative scale factor of 2) are stored in a database 10, where example images LE_i are of low-resolution and example images HE_i are their high resolution version. It is also common to store examples which have a scale gap which is the same as the scale gap between the desired high resolution image and the input low-resolution image. For each patch PT in the input low resolution image L, a similar patch LE(PT) is found in one of the low-resolution example images LE_i in database 10. A corresponding high-resolution patch HE(PT) is then copied to the unknown high-res image H at the appropriate location.

[0013] Example-based super-resolution thus utilizes the known correspondences between low resolution and high resolution image patches in database 10. Higher super-resolution factors have often been obtained by repeated applications of this process.

[0014] It is important to distinguish between image interpolation (or image zooming) and super-resolution. In image interpolation, the goal is to magnify (scale-up) an image while maintaining the sharpness of the edges and the details in the image. In contrast, in super-resolution the goal is to recover new (i.e. missing) high-resolution details beyond the Nyquist frequency of the low-resolution image that are not explicitly found in any individual low-resolution image. In classical super-resolution, this high-frequency information is assumed to be split across multiple low-resolution images, implicitly found there in aliased form. In example-based super-resolution, this missing high-resolution information is assumed to be available in the high-resolution database patches, and is learned from the low-resolution/high-resolution pairs of examples in the database.

SUMMARY OF THE INVENTION

[0015] There is therefore provided, in accordance with an embodiment of the invention, a method implementable on a computing device. The method includes exploiting data redundancy to combine high frequency information from at least two different scales of an input signal to generate a super resolution version of the input signal.

[0016] There is also provided, in accordance with an embodiment of the invention, a method implementable on a computing device which includes exploiting recurrence of data from an input signal in at least two different scales of at least one reference signal to extract and to combine high frequency information from a plurality of scales of the at least one reference signal to generate a super resolution version of the input signal.

[0017] Moreover, in accordance with an embodiment of the invention, the at least one reference signal is the input signal.

[0018] Further, in accordance with an embodiment of the invention, the at least one reference signal may be either the input or a transformed version of the input signal. The transformations may be rotation, reflection, downscaling, upscaling, similarity transformation, affine transformation, projective transformation, a color transformation, an intensity transformation, any geometric transformation or any function applied to the input signal.

[0019] Still further, in accordance with an embodiment of the invention, the at least one reference signal may be an intermediate resolution signal obtained as an intermediate result from the input signal.

[0020] Alternatively, in accordance with an embodiment of the invention, the signals may be an image, a video sequence, an fMRI sequence, an audio signal, an ultrasound signal, MRI signal, CT signal, radar signal, range sensor signal, electrical signal, medical signal or an N-dimensional signal (where N=1,2,3,4,...).

[0021] Moreover, in accordance with an embodiment of the invention, the exploiting includes using patches to find repetitive patch information, such as patch colors, patch intensities, patch Laplacians, high-pass filtered versions of the patch, low-pass filtered versions of the patch, band-pass filtered versions of the patch, patch descriptors or any function applied to the patch.

[0022] Further, in accordance with an embodiment of the invention, the patches are smaller in at least one dimension than the input signal. For example, the patches may be image patches, spatio-temporal video patches, spatial video patches, temporal video patches or n-dimensional signal patches in an N-dimensional signal (where N=1,2,3,4,... and $n \leq N$).

[0023] Still further, in accordance with an embodiment of the invention, each of the scales may be a spatial scale, a temporal scale, a scale along a dimension in an N-dimensional signal, a color scale, an intensity scale, a dynamic range scale or a spectral range scale.

[0024] Additionally, in accordance with an embodiment of the invention, the exploiting utilizes data recurrence within the input scale of the input signal and across at least one scale downscale from the input scale.

[0025] Moreover, in accordance with an embodiment of the invention, at least one of the different scales has a scale gap relative to the input scale of the input signal which is different than the scale gap between the input scale and the output scale of the super-resolution version of the input signal. The at least one of the different scales may have a non-dyadic scale gap relative to the input scale of the input signal.

[0026] Further, in accordance with an embodiment of the invention, the exploiting includes for each input patch in at least a part of the input signal, finding at least one matching patch within a plurality of scaled versions of the input signal and solving for a high-resolution signal. The solving includes recovering intermediate-resolution details of at least

one intermediate up-scaled version of the input signal, the details being extracted from at least one higher-scale parent patch of at least one of the matching patches and bridging the gap between the intermediate-resolution scale and a high-resolution scale of the high-resolution signal.

[0027] Still further, in accordance with an embodiment of the invention, the method includes transforming either the patch or at least one matching patch as part of finding similarity of patch information. Such a transformation might be by rotation, translation, reflection, scaling, similarity, affine and projective, color transformation or intensity transformation.

[0028] Moreover, in accordance with an embodiment of the invention, the details might be from the parent patch, a Laplacian of the parent patch, a high-pass filtered version of the parent patch, a weighted average of a plurality of parent patches, a weighted average of a plurality of filtered versions of parent patches, colors of the parent patch, intensities of the parent patch or a band-pass filtered version of the parent patch.

[0029] Additionally, in accordance with an embodiment of the invention, the bridging includes solving equations relating the high-resolution scale to the intermediate-resolution scale. The solving may include weighting each of the equations according to the degree of similarity of the input patch to the matching patch.

[0030] Moreover, in accordance with an embodiment of the invention, the high-resolution signal may be the super-resolution version of the input signal or an intermediate-resolution version of the input signal between the scale of the input signal and the scale of the super-resolution version of the input signal.

[0031] Further, in accordance with an embodiment of the invention, the exploiting includes coarse-to-fine repetitions of the finding and solving. The scale of the high-resolution signal ranges from the input scale of the input signal to the scale of the super-resolution version of the input signal. At each repetition, the high-resolution signal may be added to the at least one reference signal. In one case, the intermediate-resolution scale may be the high-resolution scale.

[0032] Still further, in accordance with an embodiment of the invention, the equations are a function of a blur kernel associated with the scale gap between the high-resolution scale and the intermediate scale. The equations may also be a function of a residual sub-unit

misalignment in at least one dimension between the input patch and the at least one matching patch.

[0033] Moreover, in accordance with an embodiment of the invention, the blur kernel may be a point spread function of a sensor, a Gaussian blur, a low-pass filter or an exposure-time function of a video camera. Alternatively, in accordance with an embodiment of the invention, the blur kernel may be estimated from matched patches across scales.

[0034] Still further, in accordance with an embodiment of the invention, the super-resolution provides an increase in spatial resolution, temporal resolution, color resolution, intensity resolution, dynamic range, spectral range or at least one dimension of the input signal.

[0035] Additionally, in accordance with an embodiment of the invention, the super-resolution version of the input signal may be used for signal up-sampling, digital zoom-in, viewing in slow-motion, increasing dynamic-range, changing aspect ratio, changing frame-rate, as part of digital-to-analog conversion, image compression, signal compression, video compression, guided signal acquisition to maximize resolution after processing, guided acquisition of medical signals to decrease radiation, signal format conversion, extracting new samples of the input signal, reducing blur or reducing aliasing.

[0036] There is also provided, in accordance with an embodiment of the invention, a method implementable on a computing device which includes generating a super resolution version of a single input video sequence in at least the temporal dimension.

[0037] Further, in accordance with an embodiment of the invention, the generating includes adding new video frames to reduce at least one of: motion aliasing and motion blur.

[0038] Still further, in accordance with an embodiment of the invention, the generating includes exploiting recurrence of data from the input video sequence in at least one scale of at least one reference video sequence to recover at least high temporal frequency information from the at least one reference video sequence. The at least one reference video sequence is not obtained by a different simultaneous recording of the same scene. It may be the input video sequence and it may be a transformed version of the input video sequence, such as by rotation, reflection, downscaling, upscaling, similarity transformation, affine transformation, projective transformation, a color transformation, an intensity transformation, any geometric transformation or a function applied to the input video sequence.

[0039] Moreover, in accordance with an embodiment of the invention, the at least one reference video sequence may be an intermediate resolution video sequence generated from the input video sequence.

[0040] Furthermore, in accordance with an embodiment of the invention, the exploiting includes using patches to find repetitive patch information, such as patch colors, patch intensities, patch Laplacians, filtered versions of the patch, patch descriptors or any function applied to the patch.

[0041] Still further, in accordance with an embodiment of the invention, the at least one scale may be at least one of: a spatial scale, a temporal scale, an input scale of the input video sequence, a color scale, an intensity scale, a dynamic range scale, a spectral range scale, and a non-dyadic scale gap relative to the input scale in at least one of space and time.

[0042] Moreover, in accordance with an embodiment of the invention, the exploiting includes for each input patch in at least a part of the input video sequence, finding at least one matching patch within the at least one reference video sequence or the scaled version of the at least one reference video sequence and recovering at least high temporal frequency information from the matching patches to generate the super resolution version of the input video sequence.

[0043] Further, in accordance with an embodiment of the invention, the recovering includes extracting at least high temporal frequency information from at least one higher scale parent patch of the at least one matching patch.

[0044] Moreover, in accordance with an embodiment of the invention, the recovering includes finding the at least one matching patch at sub-frame misalignment accuracy and solving equations based on the sub-frame misalignment and/or a blur function relating a high resolution scale to the scale of at least one of the input patch and the matching patch.

[0045] Additionally, in accordance with an embodiment of the invention, the method also includes transforming the patch and/or the at least one matching patch as part of finding similarity of patch information. For example, the transformation may be rotation, translation, scaling, reflection, similarity, affine, projective, color transformation or intensity transformation.

[0046] Moreover, in accordance with an embodiment of the invention, the super-resolution version of the input video sequence may be used for frame rate conversion, as part

of digital to analog conversion, video compression, temporal zoom-in, guided video acquisition to maximize resolution after processing, extracting new intermediate frames of the input video, reducing blur, reducing temporal aliasing or reducing motion aliasing. For example, the super resolution version may be a slow motion version or an up-sampled version.

[0047] Still further, in accordance with an embodiment of the invention, the finding includes performing an approximate nearest neighbor search.

[0048] There is also provided, in accordance with an embodiment of the invention, a method implementable on a computing device which includes obtaining at least temporal super-resolution of an input video sequence using recurrence of data from the input video sequence found within an external database of examples of low-resolution / high-resolution pairs of video sequences.

[0049] There is also provided, in accordance with an embodiment of the invention, apparatus including a receiver to receive an input signal and a super resolution generator to exploit data redundancy in the input signal and to combine high frequency information from at least two different scales of the input signal to generate a super resolution version of the input signal.

[0050] There is still further provided, in accordance with an embodiment of the invention, apparatus including a receiver to receive an input signal and a super resolution generator to exploit recurrence of data from the input signal in at least two different scales of at least one reference signal to extract and to combine high frequency information from a plurality of scales of the at least one reference signal to generate a super resolution version of the input signal.

[0051] There is still further provided, in accordance with an embodiment of the invention, apparatus including a receiver to receive a single input video sequence and a super resolution generator to generate a super resolution version of the single input video sequence in at least the temporal dimension.

[0052] Moreover, in accordance with an embodiment of the invention, the generator exploits recurrence of data from said single input video sequence in at least one scale of at least one of: said single input video sequence and at least one reference video sequence which is not a different simultaneous recording of the same scene.

[0053] Finally, there is provided, in accordance with an embodiment of the invention, a method implementable on a computing device which includes generating a super resolution version of a single input audio signal by exploiting recurrence of data from the single input audio signal in at least one scale of either the single input audio signal or at least one reference audio signal to recover at least high frequency information.

BRIEF DESCRIPTION OF THE DRAWINGS

[0054] The subject matter regarded as the invention is particularly pointed out and distinctly claimed in the concluding portion of the specification. The invention, however, both as to organization and method of operation, together with objects, features, and advantages thereof, may best be understood by reference to the following detailed description when read with the accompanying drawings in which:

[0055] Fig. 1 is a schematic illustration of prior art super resolution methods;

[0056] Fig. 2 is a set of three images, each image being a smaller version of its neighbor, showing data recurrence between the images;

[0057] Fig. 3 is a graph of the extent of data recurrence in a standard image database;

[0058] Figs. 4A and 4B together are schematic illustrations of a super resolution method, constructed and operative in accordance with an embodiment of the invention;

[0059] Fig. 5 is a set of four images, an input, low resolution image, a bi-cubic interpolation version, its super resolution version and the original image from which the low resolution image was taken; and

[0060] Fig. 6 is a schematic illustration of different images and different scale factors as used in the invention.

[0061] It will be appreciated that for simplicity and clarity of illustration, elements shown in the figures have not necessarily been drawn to scale. For example, the dimensions of some of the elements may be exaggerated relative to other elements for clarity. Further, where considered appropriate, reference numerals may be repeated among the figures to indicate corresponding or analogous elements.

DETAILED DESCRIPTION OF THE INVENTION

[0062] In the following detailed description, numerous specific details are set forth in order to provide a thorough understanding of the invention. However, it will be understood by those skilled in the art that the invention may be practiced without these specific details. In other instances, well-known methods, procedures, and components have not been described in detail so as not to obscure the invention.

[0063] Applicants have realized that single images contain high resolution information in them and that there is no need to refer to an external database for the high resolution information. Applicants have observed that small patches repeat many times in the images, both within scale (i.e. the same image scale as the input image) and across scale (i.e. from scales other than that of the input image). This redundancy may be exploited for all types of visual items.

[0064] For simplicity of discussion, the disclosure begins by describing the method on images. However, it will be appreciated that the invention may be implemented on all types of visual items as well as other types of signals.

[0065] Reference is now made to Fig. 2, which shows an input image 20 and two lower resolution versions of it, images 22 and 24. Applicants have realized that small patches repeat themselves elsewhere in an image, often at subpixel misalignments. For example, in image 20, various patches 26 of the roof of the building are very similar and various patches 28 of the field are also very similar. Patch recurrence within the same image scale may be utilized for classical super-resolution, particularly since the similarity may be at sub-pixel shifts.

[0066] However, Applicants have also realized that patches repeat not only within the same image scale, but also across different scales. For example, patch 30A repeats as patch 30B in lower resolution image 24 and patch 32A matches patch 32B in lower resolution image 22. Such repetitions may also be found under a transformation (e.g., rotation, reflection, affine, etc.).

[0067] The B patches 30B and 32B match the A patches, but correspond to other parts of input image 20. Thus, patch 30C which is a high resolution version of patch 30B may also serve as a high resolution version of 30A. Similarly, patch 32C is the high resolution version of patch 32B and may be used as a high resolution version of patch 32A. The B and C patches may now provide high resolution and low resolution pairs of patches which may be

utilized in example-based super-resolution to determine the super resolution version of the A patches.

[0068] It will be appreciated that, with the within scale and across scale patch similarities, the invention may create a high resolution image from a single image, without using any external information, such as an external database or prior examples. As will also be discussed hereinbelow, it may also be applicable to spatial or temporal resolution enhancement of video sequences, of single dimensional signals, such as audio or speech, and of multi-dimensional signals.

[0069] Reference is now briefly made to Fig. 3, which illustrates how much patch redundancy there is within natural images. Over hundreds of natural images in the Berkeley Segmentation Database, the number of patches which have at least 10 similar patches (the notion of similarity is discussed below) is over 70%, even in a significantly reduced image of 0.25 of the original image. This holds true even when areas of uniform color (such as a blue sky) are removed. In the original image, about 90% of the patches have at least 10 similar patches. Thus, as argued above, the repetition of patches within and across scales may be utilized to generate a super-resolution version of the original image.

[0070] The patches may be of any size, but small image patches (e.g., 5 x 5 pixels) in a natural image tend to redundantly recur many times inside the image. For example, 5x5 is the patch size used to produce the statistics in Fig 3.

[0071] Patch similarity may be measured by defining a distance between patches. For example, a Gaussian-weighted SSD (sum-of-squares distance) may be utilized. Note that textured patches tend to have much larger SSD errors than smooth (low-variance patches) when compared to other very similar-looking patches (especially in the presence of inevitable sub-pixel misalignments). To neutralize this, a distance threshold may be defined for each patch by measuring the Gaussian-weighted SSD distance w of the patch to a slightly misaligned copy of itself (by 0.5 pixel). For the statistics shown in Fig. 3, patches with a distance w below this threshold are considered similar to the source patch.

[0072] When searching for patches, it is possible to search amongst patches as is or to search in a manner which is invariant to different properties which may not be important for the high resolution information. Thus, for example, the method may search amongst patches which have had their mean values subtracted, thus achieving invariance to the mean color of

the patch. Any other kind of low pass filtering may be applied to the patches before searching for matches.

[0073] Reference is now made to Figs. 4A and 4B which, together, illustrate a novel super-resolution method, constructed and operative in accordance with an embodiment of the invention. Fig. 4A illustrates a within scale part of the method utilizing patches from the same scale and Fig. 4B illustrates an across scale part utilizing patches from different scales.

[0074] In Fig. 4A, an exemplary calculation for a main pixel p is shown. First, patches similar to patch P around pixel p are found within the input, low-resolution image L. Two exemplary patches P_1 and P_2 are found, around pixels p_1 and p_2. Each one of these pixels provides a linear equation constraining the unknown pixel values in the high resolution image

[0075] As can be seen, pixel p and each pixel p_i induce one linear constraint on the unknown high-resolution intensity values within a local neighborhood N or N_i respectively around high-resolution pixel q (which is the pixel in the high resolution image corresponding to pixel p), where the size of the local neighborhood is determined by the "support" of the blur kernel B. Equation 2 (repeated here) defines the linear constraint:

[0076] $$L(p) = (H * B_j)(q) = \sum_{q_i \in Support(B_j)} H(q_i) B_j(q_i - q) \qquad \text{Equation 2}$$

where $H(q_i)$ are the unknown high-resolution intensity values and B_j is the same blur kernel B relating high resolution image H with low resolution L as before, but shifted according to the subpixel misalignment between patches P and P_j, determined during the match process.

[0077] Fig. 4B illustrates the across-scale part of the method of the invention. For this part, it is possible to imagine a plurality of unknown upscale images (images I_1 and I_2 are shown in Fig. 4B but many more are possible and are part of the invention) between input, low resolution image L (equivalent to image I_0) and output super-resolution image H. These upscale images have corresponding down-scale images I_{-1} and I_{-2} of input image L. The scaling factors may be any scaling factor, such as 1.25, 2, 3, etc. The different scaling factors can be chosen arbitrarily, or they can be uniformly spaced e.g. $2^{(1/3)}$, $2^{(2/3)}$, 2.

[0078] Each upscale image has its own blur kernel B^{Ij}, which may be more compactly supported (i.e. have a smaller spread), than the point spread function for input image L. It will be appreciated that the blur kernels B^{Ij} for the across-scale part of the method compensate only for the residual gap in scale between the resolution level of the upscale image and the

final resolution level of the target high-resolution image H. The blur kernels B^{Ij} may be approximated as a Gaussian function.

[0079] The across-scale part of the method may search for matching patches in down-scale images I_1 and I_2 and may select the patch which matches best or may select any desired number k (k = 1,2,3,...) of good matches. Fig. 4B shows an example patch $P(I_{-2})$ which matches initial patch P, but was found in down-scale image I_{-2}. Matching patch $P(I_{-2})$ is a down-scale version of patch $LP(I_{-2})$ of input image L. Thus, matching patch $P(I_{-2})$ and its corresponding patch $LP(I_{-2})$ are a low-resolution/high-resolution pair of images, generated from single input image L. Accordingly, corresponding patch $LP(I_{-2})$ may be copied to upscale image I_2 at the location corresponding to the initial point p, as $LP(I_{-2})(p)$, in an operation similar to that of example-based super-resolution methods. Corresponding patch $LP(I_{-2})$ may be considered a "learned" patch.

[0080] However, upscale image I_2 is not the desired high resolution image H and thus, the method of the invention may solve for the residual scale gap, by implementing Equation 2 and using the B^{I_2} blur kernel, which kernel has a smaller spread than $B = B_L$, the kernel for input image L.

[0081] The method of the invention may search all of the down-scale versions for matching patches using any appropriate nearest neighbor algorithm. Such an algorithm may be exact or may be approximate. Exemplary methods include the Approximate Nearest Neighbor search such as the article by S. Arya and D.M. Mount, "Approximate nearest neighbor queries in finite dimensions", in *SODA,* 1993, or the PatchMatch nearest neighbor search, such as described in the article by Connelly Barnes, Eli Shechtman, Adam Finkelstein, and Dan B Goldman entitled "PatchMatch: A Randomized Correspondence Algorithm for Structural Image Editing", *ACM Transactions on Graphics (Proc. SIGGRAPH)*, Vol. 28(3), Aug 2009.

[0082] A patch may be regarded as a very small image. Thus, methods for image-to-image alignment may be used to align image patches at sub-pixel accuracy (e.g., using the direct gradient-based methods described in the following articles: M. Irani, B. Rousso, and S. Peleg, "Computing Occluding and Transparent Motions", *International Journal of Computer Vision (IJCV)*, Vol. 12, No. 1, pp. 5-16, Feb 1994, and M. Irani and S. Peleg, "Motion Analysis for Image Enhancement: Resolution, Occlusion, and Transparency", *Journal of Visual*

Communication and Image Representation, Vol. 4, No. 4, pp. 324-335, Dec 1993, which methods bring images into alignment at high sub-pixel accuracy).

[0083] Another possible method to obtain patch matches at sub-pixel accuracy is to extract all the patches from the input image (input signal) at pre-determined sub-pixel (sub-unit) shifts, using interpolation. These extracted patches may serve as reference patches for matching to the original image (signal) patches, thus obtaining matches at sub-pixel (sub-unit) matches.

[0084] The method may then copy the matched patch to the associated upscale images and may solve for the residual scale gap (i.e. generate equation 2) from the upscale image, such as I_2 or I_1, to high resolution image H. If the downscale image is scaled down by the amount that image H is scaled up from input image L, then there is no residual scale gap and the matched patch may be copied directly to high resolution image H.

[0085] Alternatively, the method may utilize a multiple k of nearest neighbors, rather than the closest neighbor. For example, k may be 9.

[0086] It will be appreciated that, when searching for similar patches for example-based super-resolution, any measure of similarity may be utilized, and that the patches may be measured to be similar as is or under a parametric transformation, such as a rotation, a scale, a reflection, an affine rotation, similarity, affine, etc.

[0087] In one embodiment, the method may generate the multiple scales from the input image and may then search for input patches under various transformations in the various scales. Alternatively, it may transform the image using a range of transformations. For each transformed input image, it may generate a set of upscale/downscale images and may search for the input patches 'as is' in those scales.

[0088] It is also possible to search for matches in other transformations of the input images which are not necessarily geometric. Transformations can be applied to the color or intensity of the image or to any filtered version of the image.

[0089] Patch similarity between patch P and patch P_i may be computed using different similarity measures. One option is to use exp(beta*SSD) where SSD is the sum of squared differences between patch P_i and patch P and beta is a negative scalar. A Gaussian weighted SSD may also be used.

[0090] The process described hereinabove, when repeated for all pixels in L, will yield a large collection of (possibly overlapping) suggested learned, intermediate-resolution patches LP(p) with a range of resolution levels between input image L and output image H. Each such "learned" patch LP(p) induces linear constraints on the unknown target image H.

[0091] It will be appreciated that the matches come from both within scale and across scales. This results in a large set of linear equations which may be solved simultaneously. It is also possible to use each component separately, i.e. rely only on within scale matches or only on across scale matches.

[0092] For increased numerical stability, each equation induced by a patch P_i may be globally scaled by the degree of similarity of P_i to its source patch P. Thus, patches of higher similarity to P will have a stronger influence on the recovered high-resolution image than patches of lower similarity. The process may then utilize standard linear equation solvers to solve a weighted linear set of equations to recover super resolution image H.

[0093] It will be appreciated that the closer the learned patches are to the resolution of target image H, the better conditioned the resulting set of equations may be.

[0094] If desired, the set of linear equations may be solved gradually, coarse-to-fine. Thus, the upscale (or intermediate) images may be recovered one at a time, from the lowest to the highest resolution. When solving the equations for an intermediate image, the process may utilize not only the low-resolution/high-resolution patch correspondences found in the associated low-resolution image, but also all of the newly learned patch correspondences from any previously recovered upscale images. The process may be repeated until the resolution level of target image H may be achieved.

[0095] A simplified reconstruction can also be performed without solving equations. Each intermediate level may be reconstructed by taking a weighted average of the overlapping example patches according to their degree of patch similarity. The weighting may also be influenced by the location of each pixel relative to the patch center.

[0096] Note that in the case when all the "learned" parent patches are copied to the highest resolution level, then the set of linear equations reduces to the trivial identity equations (or just to simple weighted averaging of overlapping parent patches). In such a case, there is no need to solve any equations in order to recover the high-resolution image, as it comes out trivially.

[0097] To further guarantee consistency of the recovered high-resolution results, when a new upscale image may be obtained, it may be projected back towards input image L (such as by blurring and subsampling) and then compared to input image L. Any large differences may indicate errors in the corresponding high-resolution pixels and may be "back-projected", as described in the article by Michal Irani and Shmuel Peleg, "Improving resolution by image registration", *CVGIP: Graphical Model and Image Processing*, Vol. 53,(3), 1991, pp 231-239, onto the upscale image to correct the errors. This process may verify that each newly recovered upscale image I_1, I_2 or H may remain consistent with input low resolution image L.

[0098] A simplified version of the coarse to fine calculation may be performed as follows. Sequentially, each intermediate image between the input resolution and the desired high resolution scale may be reconstructed by taking a weighted average of the overlapping learned patches according to their degree of patch similarity. The weighting may also be influenced by the location of each pixel relative to the patch center. Once the intermediate level is reconstructed, it may or may not be back-projected to the scale of the input image to correct errors. It may then be added to the database of example patches. Subsequent reconstructions may utilize it as examples.

[0099] Reference is now made to Fig. 5, which illustrates an exemplary result of implementing the present method on an image of an eye chart, shown as "(d) ground truth image". The input image (a) was a down-scaled (blurred and subsampled) version of (d), down-scaled by a factor of 3. Image (b) shows the result of a bicubic interpolation of the input image (a), a standard method of generating an upscaled image from a single input image. Image (c) shows the result of the present method.

[00100] Note the bottom part of image (c). The lines of letters have been recovered quite well, especially when compared with image (b), due to the existence of cross-scale patch recurrence in those image areas. However, the small digits on the left margin of the image could not be recovered, since their patch recurrence occurs only within the same (input) scale.

[00101] It will be appreciated that the main improvement in resolution comes from the addition of the example-based component to the super resolution operation. However, the classical-super resolution component plays a central role in preventing the example-based super resolution component from hallucinating erroneous high-resolution details (a problem alluded to by the article by Freeman, William T., Jones, Thouis R. and Pasztor, Egon C.

entitled "Example-Based Super-Resolution" in *IEEE Comput. Graph. Appl.*, Vol 22, 2002, pp 56 – 65).

[00102] It will be appreciated that the super resolution method of the invention may be equivalently posed as optimizing an objective function with a "data-term" and two types of "prior-terms": The data-term may stem from the blur and subsample relation of the classical super resolution between high-resolution image H and low-resolution image L. The example-based super resolution constraints form one type of prior term and the multiple patches of the classical super resolution constraints form another type of prior term (at sub-pixel accuracy). The high-resolution image H which optimizes this objective function must satisfy both sets of constraints simultaneously. Similarly, a simpler objective function may be posed and optimized when using only a subset of the above-mentioned priors (e.g., only the data term + the example-based prior).

[00103] It will be appreciated that the invention may be applied in other contexts of super resolution. For example, in accordance with an embodiment of the invention, example-based cross-scale constraints may be added to classical super resolution of multiple low-resolution images of the same scene. Similarly, existing example-based super resolution methods which work with an external database may be extended by adding the unified super resolution constraints described hereinabove.

[00104] Moreover, this approach may be applied to provide super resolution of video data. It may be used for spatial super resolution of video frames (i.e., improving the spatial resolution of video frames by increasing the spatial sampling rate), and/or for temporal super resolution of the video data (i.e., increasing the frame-rate of the video sequence by generating new in-between frames. This may be done, e.g., by using recurrence of space-time patches within the video data, within and/or across spatial and/or temporal scales). This, again, may be done with or without additional external data, or additional video sequences. The result may be super-resolution from a single video sequence.

[00105] It will also be appreciated that the invention may be modified for spatio-temporal super resolution of video data. This embodiment may use a database of low-resolution / high-resolution pairs of space-time patches to increase the spatial resolution, the temporal resolution or both of them for the input video sequence. For example, this embodiment of the method may search for a good match to a query patch amongst the low-resolution patches in

the database. The corresponding high resolution patch may provide information about the form of the high resolution video at the query location. In an embodiment, the low-resolution / high-resolution space-time patches may be taken from any video sequence or sequences and may be stored in the database. Moreover, additional pairs may be added to the database by extracting low-resolution / high-resolution patches from the input sequence and scaled versions of the input sequence where the scaling may be in space in time or in both.

[00106] The space-time patches may have different dimensions along x, y, and t (time). They may be small space-time patches (e.g., 5x5x5), or space-time chunks of a few full consecutive video frames (i.e., space-time patches of **N**x**M**x**t**), or may have only a temporal component (i.e., depicting the temporal behavior of a single pixel over time, which corresponds to a space-time patch of **1**x**1**x**t**), etc. A space-time patch may be regarded as a small video sequence. Thus, methods for sequence-to-sequence alignment may be used to align space-time patches (e.g., using the method of US 7,428,345, which is incorporated by reference, which brings video sequences into alignment in space and in time at high sub-pixel and sub-frame accuracy).

[00107] It will be appreciated that the method of the invention may be utilized to create a slow motion sequence which displays events that occurred faster than the original frame-rate, and thus are not visible in the original video. This may eliminate motion aliasing, such as the Wagon Wheel Effect. Temporal super resolution may enable new higher temporal frequencies (beyond the limit defined by the Nyquist frequency of the original frame-rate) to be recovered. This is possible by exploiting internal temporal redundancy at sub-frame displacements within the same temporal scale, or by 'learning' the extra missing high-frequency information by exploiting space-time redundancy across other temporal scales of the same input video.

[00108] Temporal super-resolution may also be utilized to remove temporal aliasing (such as motion aliasing, as discussed hereinabove, or intensity aliasing, and/or for decreasing motion blur. In the invention, the reduction of motion blur occurs without having to know or estimate the motion of the moving objects.

[00109] The invention may also be utilized to improve digital to analog video converters (DACs). Existing DACs use various interpolation methods in order to generate the continuous output signal. However, they do not recover the lost high frequencies which, as

described hereinabove, are recoverable with the invention. An improved video DAC may implement the super resolution method described hereinabove to increase the resolution of the digital signal prior to creating the continuous signal, thereby adding denser sampling points while recovering missing high temporal frequencies.

[00110] The invention may also transform a video sequence from one frame-rate to another, such as is necessary when changing formats (e.g., PAL to NTSC, and vice versa), or when changing devices/displays, some of which have a specific frame-rate (various TVs, mobile phone, computer screens, digital screens, analog screens, etc.).

[00111] The aspect ratio of an image or a video sequence may be changed by increasing the sampling rate more in one dimension than in the other and/or by looking for example patches in lower scales with different x-scale factors, y-scale factors, and/or t-scale factors.

[00112] It will be appreciated that the invention may be applied to 1-D signals, 2-D signal, 3-D signals and N-dimensional signals for $N > 3$. It may also be applied to different kinds of signals such as audio, general visual data, video, images, fMRI, ultrasound, medical data, CT, MRI, PET, radar, biological data, etc. The different scales may be spatial scales, or temporal scales, or different scales along any dimension in an N-dimensional signal.

[00113] Similarly, the increase in resolution may be along any of these dimensions: spatial super resolution, temporal super resolution, or super resolution in any dimension in an N-dimensional signal. The patches searched may be either spatial image patches of spatio-temporal video patches, or N-dimensional signal patches.

[00114] The above algorithm and the patches used may be applied within a variety of different data representations. These include (but are not limited to) scaled versions of the same image, a Gaussian pyramid (but with arbitrary scale factors, and not necessarily by a factor of 2), a Laplacian pyramid (arbitrary scale factors), Gradient pyramid, first and/or second directional derivatives, or any other descriptor types (e.g., SIFT, etc.).

[00115] The scale may be changed along any appropriate direction present in the input signal. In an image, the scales may be spatial. However, for a video sequence, the scales may be spatial or temporal or spatio-temporal. Similarly for N-dimensional signals.

[00116] The scale gap between successive scales (up and down) may be uniform. Alternatively, there may be general and arbitrary scale ratios relative to the input image. These ratios may be dyadic (powers of 2) or non-dyadic.

[00117] The above algorithm may be applied to color data in a variety of possible ways. In one embodiment, the super resolution operation may be performed on the intensity component only (i.e. the image may first be transformed from the standard RGB representation to the YIQ representation) and the super resolution operation may be applied to the Y (intensity) channel only. The method may then interpolate (such as via a bi-cubic interpolation) the chromatic and illumination components (which are generally characterized by low frequency information) and may then recombine the results.

[00118] In another embodiment, the super resolution operation may be performed on each color band separately (in any color representation: RGB, YIQ, Lab, etc) and then the results may be combined.

[00119] In a further embodiment, the super resolution operation may be performed on the combined color patches (all 3 bands simultaneously), i.e., searching for nearest-neighbors on color patches.

[00120] In a still further embodiment, the nearest-neighbor search in the illumination component may guide the super resolution in all three color components. Thus, the method may first search using the Y channel patches and may use the indices of the resulting matches to perform super resolution on the other color channels.

[00121] In a further embodiment, super resolution may be utilized to obtain high dynamic range, high-resolution of color information. In this embodiment, the patch search may be implemented in multiple intensity/color scales. Different scales in the intensity/color domain correspond to different degrees of intensity/color quantization. Exploiting low-res/high-res pairs of patches obtained from different intensity/color scales provides examples for recovering high-resolution information in the intensity/color domain. In addition, patches may be combined that are offset by fractions of color units or fractions of intensity units. The result is an extension of the dynamic range of a single image by exploiting internal data redundancy within and across color/intensity scales.

[00122] This idea may be applied to a single input image using internal data redundancy, or exploiting external data redundancy relative to an external database of examples of images of a different scene.

[00123] In a further embodiment, the blur function may be estimated from the data itself, by computing the mapping between repetitive patterns of different scale within the data. For

example, if a patch P_0 in input image L has a similar patch $P(I_{-2})$ in down-scale image I_{-2}, then a more accurate blur function relating these two scales may be estimated directly from the relation between the high-resolution parent $LP(I_2)$ and patch P_0.

[00124] It will be appreciated that the invention may generate super-resolution by combining high-resolution information from two or more scales of at least one image. The high resolution information may include high-frequency information, which typically increases the quality of the image.

[00125] The method may exploit internal data redundancy across at least two scales, which may be the input scale ("within scale"), or any other scales of the image ("across scale"). For examples, the two scales may be the input scale and another scale (using similarity "within and across scales"), or just two different low-resolution scales extracted from the image. The method may be implemented with or without the classical super-resolution constraints that exploit redundancy within the same scale and with or without solving sets of equations.

[00126] Combining high-frequency information obtained from two or more different scale gaps allows recovering high-frequency information that is not fully contained in any of the individual scale-gaps. This happens in cases when the vanishing frequencies of the blur functions relating these different scale gaps, are different. This is particularly important in the case of temporal super resolution in video, where the blur function is the exposure time of the video camera, which is shaped like a Rect function. Its Fourier transform is a Sinc function, which has multiple vanishing frequencies. Different exposure times (different Rects) will have different Sinc functions. If there are scale gaps whose Rects have different well-separated vanishing frequencies, new high frequency information may be recovered that is not fully contained in any of the individual scales.

[00127] This may be possible with different scales of the input image (as discussed hereinabove), or with multiple scales from within an external database image. Alternatively, classical multi-frame super resolution methods may be extended to use also example-based information obtained from other scales of the input images. This may add more equations or information learned from self-similarities across scales within the input images (or within an external database of images).

[00128] The high-frequency information may be combined from the different scales with or without solving equations. It may be combined by solving linear sets of equations (as

described hereinabove), or by a simple weighted average of the information of the patches obtained from the different scales (e.g., by averaging the patch information, or by averaging only the high-frequency details of the patches, or by incorporating these into the Markov Random Field used by Freeman et al., etc.). The weights may be according to any appropriate measure, such as according to the degree of similarity of the Nearest-Neighbor patches found, according to pixel position within the patch, etc.)

[00129] It will be appreciated that the invention may utilize examples from at least one scale gap relative to the input image which is different from the scale gap S between the input and the output images. This enables the invention to find similarities in other image scales which are not found in the 1/S image scale. An example of this may be seen in Fig. 6, to which reference is now made.

[00130] Fig. 6 shows an image of a large house 100 and a small house 102, about 1/3 the size of large house 100. Each house has one window W and one door D. The door of small house 102 is labeled D102 and the window of large house 100 is labeled W100.

[00131] When magnifying the image by a factor of 2, previous methods will look for a match for the small house (or its smaller patch components) in 1/2 of the original scale. However, none of the houses in that image scale match properly to the 1/3 sized house 102 and, thus the patch matching will fail at that level. The correct match for the small house can be found only at 1/3 of the image scale (the image labeled 1/3x), where it will match a smaller version of the larger house (i.e. its smaller patches will find a good match). The corresponding high-resolution examples will be extracted from the 2x1/3 = 2/3 scale and copied to the output at the appropriate location.

[00132] Simple interpolation will fail because the fine texture of the window and door of the small house in the input image are beyond the limit defined by the Nyquist frequency of the input sampling.

[00133] In an alternative embodiment of the invention, only the extra high-frequency details of a "learned" patch at the relevant frequency bands may be added. For this, any high pass filter, such as a Laplacian of the patch, may be performed on the patch before adding the patch into its upscale image. Another exemplary high pass filter may be to add only the difference between the learned patch and the low-resolution patch.

[00134] To generate the different image scales, the images typically are blurred and then subsampled. In an alternative embodiment, the images may just be blurred. In this embodiment, the low-res/high-res pair may be of the same size, but one may be blurry and one may be sharp. The input image L may then be increased to the output size by simple interpolation, to generate an initial high resolution image H'. The method may then utilize the interpolated blurry patches of H' as low-resolution patches to be searched for in the low-resolution/high-resolution pairs. Once a matching low-resolution patch is found in the examples, its corresponding sharp patch (the 'learned' high-resolution patch) may be copied to the appropriate position in the output image H (or just its extra high-frequencies may be added to the blurry output image H'). As before, the patches may be copied and multiple overlapping patches may be averaged, or they may induce additional equations according to the blur function which may be further solved.

[00135] The method may also be used to perform "example-based deblurring" (spatial deblurring, motion deblurring, etc.) by generating multiple scales of increasingly blurrier images/videos (blurring in space or in time). For example, example-based motion deblurring in video sequences may be obtained even without increasing the frame-rate. Imagine a video sequence with a rotating fan, which starts rotating slowly and gradually increases its speed. The frames of the slow-rotating fan may serve as sharp examples for the frames with the blurry fast rotating fan.

[00136] In an alternative embodiment, the method of the invention may provide super resolution of audio data. It may be used to improve the temporal resolution of recorded audio tracks, by increasing the sampling rate while recovering new higher temporal frequencies. The method may utilize recurrence of temporal patches within the input scale of the audio track and/or in different temporal scales of the audio track. The method may utilize external audio data, or additional audio tracks, if desired.

[00137] Super resolution of audio tracks may be utilized for an audio digital to analog converter (DAC). The conversion may be done by using the output of a super resolution algorithm, such as the one described hereinabove or any other super resolution method, applied to the input audio track, as part of a digital to analog converter. This may be utilized for digital music players, PC-sound card, and hi-fi systems, etc.

[00138] The super resolution of audio tracks may be utilized to change the temporal sampling rate of the signal when changing from one digital device (or format) to another. This requires adding new samples according to the correct high temporal frequencies (higher than the limit defined by the Nyquist frequency of the input sampling rate) that are not recoverable by regular temporal interpolation.

[00139] In accordance with an embodiment of the invention, all of the above ideas may be used for advanced image, video, and audio compression. For example, the super resolution may be utilized to decompress an image compressed with any standard compression algorithm, thereby to recover the high-resolution image. In addition, if needed, on the decoding side and prior to transmitting, the decoder may generate the super resolution image and may compare it with the true original image. Sections of large reconstruction error may be further corrected by sending extra bits.

[00140] A similar operation may be performed for a video sequence, in space and/or in time. In particular, a lower frame-rate video may be transmitted by sampling the frames (with or without temporal blur). The missing frames may be recovered on the decoding side using temporal super resolution. In addition, if needed, on the decoding side and prior to transmitting, the decoder may generate the super resolution video and may compare it with the original video. Sections of large reconstruction error may be further corrected by sending extra bits.

[00141] In a further alternative embodiment, the super resolution method may be used in the design of protocols for budgeted medical image acquisition to minimize radiation. Consider the process of scanning a 3 dimensional body with some medical imaging device such as a CT scanner. Such a scan is performed by taking a sequence of two dimensional slices of the body. The treating doctor would like to get the highest possible resolution within each slice as well as having the inter slice spacing as small as possible. However, higher resolution means longer scanning time, which incurs higher operation costs and is uncomfortable for the patient. Furthermore, longer scanning time corresponds to higher amounts of radiation being absorbed by the patient. Therefore, there is a tradeoff between scanning time and the resolution of the acquired image.

[00142] A budgeted acquisition protocol, where the budget may be set by the maximal amount of radiation that the patient may absorb, is generally desirable. The protocol ideally

should be designed such that the resulting medical image has maximal resolution in the areas of interest. In accordance with an embodiment of the invention, the super resolution method of the invention may optimize this tradeoff and may enable the protocol to spend less time scanning areas (or directions) where there is higher potential for super resolution. For example, if it is known that the content of close slices along the slicing direction is very similar to one another, then lower resolution slices may be obtained and a super resolved version of each slice, using only data of the input image, may be generated using the data redundancy across slices.

[00143] Unless specifically stated otherwise, as apparent from the preceding discussions, it is appreciated that, throughout the specification, discussions utilizing terms such as "processing," "computing," "calculating," "determining," or the like, refer to the action and/or processes of a computer, computing system, or similar electronic computing device that manipulates and/or transforms data represented as physical, such as electronic, quantities within the computing system's registers and/or memories into other data similarly represented as physical quantities within the computing system's memories, registers or other such information storage, transmission or display devices.

[00144] Embodiments of the invention may include apparatus for performing the operations herein. This apparatus may be specially constructed for the desired purposes, or it may comprise a general-purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable storage medium, such as, but not limited to, any type of disk, including floppy disks, optical disks, magnetic-optical disks, read-only memories (ROMs), compact disc read-only memories (CD-ROMs), random access memories (RAMs), electrically programmable read-only memories (EPROMs), electrically erasable and programmable read only memories (EEPROMs), magnetic or optical cards, Flash memory, or any other type of media suitable for storing electronic instructions and capable of being coupled to a computer system bus.

[00145] The processes and displays presented herein are not inherently related to any particular computer or other apparatus. Various general-purpose systems may be used with programs in accordance with the teachings herein, or it may prove convenient to construct a more specialized apparatus to perform the desired method. The desired structure for a variety

of these systems will appear from the description below. In addition, embodiments of the invention are not described with reference to any particular programming language. It will be appreciated that a variety of programming languages may be used to implement the teachings of the invention as described herein.

[00146] While certain features of the invention have been illustrated and described herein, many modifications, substitutions, changes, and equivalents will now occur to those of ordinary skill in the art. It is, therefore, to be understood that the appended claims are intended to cover all such modifications and changes as fall within the true spirit of the invention.

CLAIMS

What is claimed is:

1. A method implementable on a computing device, the method comprising:

exploiting data redundancy to combine high frequency information from at least two different scales of an input signal to generate a super resolution version of said input signal.

2. A method implementable on a computing device, the method comprising:

exploiting recurrence of data from an input signal in at least two different scales of at least one reference signal to extract and to combine high frequency information from a plurality of scales of said at least one reference signal to generate a super resolution version of said input signal.

3. The method according to claim 2 and wherein said at least one reference signal is said input signal.

4. The method according to claim 2 and wherein said at least one reference signal is at least said input and a transformed version of said input signal obtained by at least one of the following operations applied to said input signal: rotation, reflection, downscaling, upscaling, similarity transformation, affine transformation, projective transformation, a color transformation, an intensity transformation, any geometric transformation, and any function applied to said input signal.

5. The method according to claim 2 and wherein said at least one reference signal is an intermediate resolution signal obtained as an intermediate result from said input signal.

6. The method according to claim 2 wherein said signals are at least one of the following types of signals: an image and a video sequence.

7. The method according to claim 2 wherein said signals are at least one of the following types of signals: an fMRI sequence, an audio signal, an ultrasound signal, MRI signal, CT signal, radar signal, range sensor signal, electrical signal, medical signal, and an N-dimensional signal (where N=1,2,3,4,...).

8. The method according to claim 2 wherein said exploiting comprises using patches to find repetitive patch information.

9. The method according to claim 8 and wherein said patch information is at least one of the following: patch colors, patch intensities, patch Laplacians, high-pass filtered versions of the patch, low-pass filtered versions of the patch, band-pass filtered versions of the patch, patch descriptors and any function applied to the patch.

10. The method according to claim 8 wherein said patches are smaller in at least one dimension than said input signal.

11. The method according to claim 8 wherein said patches are at least one of the following: image patches, spatio-temporal video patches, spatial video patches, temporal video patches, and n-dimensional signal patches in an N-dimensional signal (where N=1,2,3,4,... and $n \leq N$).

12. The method according to claim 2 wherein each of said at least two scales is at least one of the following: a spatial scale, a temporal scale, a scale along a dimension in an N-dimensional signal, a color scale, an intensity scale, a dynamic range scale and a spectral range scale.

13. The method according to claim 3 and wherein said exploiting utilizes data recurrence within the input scale of said input signal and across at least one scale downscale from said input scale.

14. The method according to claim 3 wherein at least one of said different scales has a scale gap relative to the input scale of said input signal which is different than the scale gap between said input scale and the output scale of said super-resolution version of said input signal.

15. The method according to claim 3 wherein at least one of said different scales has a non-dyadic scale gap relative to the input scale of said input signal.

16. The method according to claim 3 and wherein said exploiting comprises:

for each input patch in at least a part of said input signal, finding at least one matching patch within a plurality of scaled versions of said input signal; and

solving for a high-resolution signal comprising:

recovering intermediate-resolution details of at least one intermediate up-scaled version of said input signal, said details being extracted from at least one higher-scale parent patch of at least one of said matching patches; and

bridging the gap between said intermediate-resolution scale and a high-resolution scale of said high-resolution signal.

17. The method according to claim 16 and also comprising transforming at least one of said patch and said at least one matching patch as part of finding similarity of patch information.

18. The method according to claim 17 and wherein said transforming comprises performing at least one of the following transformations: rotation, translation, reflection, scaling, similarity, affine and projective, color transformation and intensity transformation.

19. The method according to claim 16 and wherein said details are at least one of: said parent patch, a Laplacian of said parent patch, a high-pass filtered version of said parent patch, a weighted average of a plurality of parent patches, a weighted average of a plurality of filtered versions of parent patches, colors of said parent patch, intensities of said parent patch and a band-pass filtered version of said parent patch.

20. The method according to claim 16 and wherein said bridging comprising solving equations relating said high-resolution scale to said intermediate-resolution scale.

21. The method according to claim 20 and comprising weighting each of said equations according to the degree of similarity of said input patch to said matching patch.

22. The method according to claim 16 and wherein said high-resolution signal is one of the following: said super-resolution version of said input signal and an intermediate-resolution version of said input signal between the scale of said input signal and the scale of said super-resolution version of said input signal.

23. The method according to claim 16 wherein said exploiting comprises coarse-to-fine repetitions of said finding and solving, wherein the scale of said high-resolution signal ranges from the input scale of said input signal to the scale of said super-resolution version of said input signal, and wherein at each said repetition, said high-resolution signal is added to said at least one reference signal.

24. The method according to claim 16 and wherein said intermediate-resolution scale is said high-resolution scale.

25. The method according to claim 20 and wherein said equations are a function of a blur kernel associated with the scale gap between said high-resolution scale and said intermediate scale.

26. The method according to claim 25 and wherein said equations are a function of a residual sub-unit misalignment in at least one dimension between said input patch and said at least one matching patch.

27. The method according to claim 25 and wherein said blur kernel is one of the following: a point spread function of a sensor, a Gaussian blur, a low-pass filter and an exposure-time function of a video camera.

28. The method according to claim 25 and also comprising estimating said blur kernel from matched patches across scales.

29. The method according to claim 2 and wherein said super-resolution provides an increase in resolution in at least one of: spatial resolution, temporal resolution, color resolution, intensity resolution, dynamic range, spectral range and at least one dimension of said input signal.

30. The method according to claim 2 and wherein said super-resolution version of said input signal is used for at least one of the following applications: signal up-sampling, digital zoom-in, viewing in slow-motion, increasing dynamic-range, changing aspect ratio, changing frame-rate, as part of digital-to-analog conversion, image compression, signal compression, video compression, guided signal acquisition to maximize resolution after processing, guided acquisition of medical signals to decrease radiation, signal format conversion, extracting new samples of said input signal, reducing blur and reducing aliasing.

31. A method implementable on a computing device, the method comprising:

generating a super resolution version of a single input video sequence in at least the temporal dimension.

32. The method according to claim 31 and wherein said generating comprises adding new video frames to reduce at least one of: motion aliasing and motion blur.

33. The method according to claim 31 and wherein said generating comprises exploiting recurrence of data from said input video sequence in at least one scale of at least one reference video sequence to recover at least high temporal frequency information from said at least one reference video sequence and wherein said at least one reference video sequence is not obtained by a different simultaneous recording of the same scene.

34. The method according to claim 33 and wherein said at least one reference video sequence is said input video sequence.

35. The method according to claim 33 and wherein said at least one reference video sequence is a transformed version of said input video sequence, obtained by at least one of the following transformations in at least one dimension of at least one of space and time: rotation, reflection, downscaling, upscaling, similarity transformation, affine transformation, projective transformation, a color transformation, an intensity transformation, any geometric transformation and a function applied to said input video sequence.

36. The method according to claim 33 and wherein said at least one reference video sequence is an intermediate resolution video sequence generated from said input video sequence.

37. The method according to claim 33 wherein said exploiting comprises using patches to find repetitive patch information.

38. The method according to claim 37 and wherein said patch information is at least one of the following: patch colors, patch intensities, patch Laplacians, filtered versions of the patch, patch descriptors, and any function applied to the patch.

39. The method according to claim 33 and wherein said at least one scale is at least one of: a spatial scale, a temporal scale, an input scale of said input video sequence, a color scale, an intensity scale, a dynamic range scale, a spectral range scale, and a non-dyadic scale gap relative to said input scale in at least one of space and time.

40. The method according to claim 33 and wherein said exploiting comprises:

for each input patch in at least a part of said input video sequence, finding at least one matching patch within at least one of: said at least one reference video sequence and said scaled version of said at least one reference video sequence; and

recovering at least high temporal frequency information from said matching patches to generate said super resolution version of said input video sequence.

41. The method according to claim 40 and wherein said recovering comprises extracting at least high temporal frequency information from at least one higher scale parent patch of said at least one matching patch.

42. The method according to claim 40 and wherein said recovering comprises finding said at least one matching patch at sub-frame misalignment accuracy and solving equations based on at least one of: said sub-frame misalignment and a blur function relating a high resolution scale to the scale of at least one of said input patch and said matching patch.

43. The method according to claim 40 and also comprising transforming at least one of said patch and said at least one matching patch as part of finding similarity of patch information.

44. The method according to claim 43 and wherein said transforming is at least one of: rotation, translation, scaling, reflection, similarity, affine, projective, color transformation, and intensity transformation.

45. The method according to claim 31 and wherein said super-resolution version of said input video sequence is used for at least one of the following applications: frame rate conversion, as part of digital to analog conversion, video compression, temporal zoom-in, guided video acquisition to maximize resolution after processing, extracting new intermediate frames of said input video, reducing blur, reducing temporal aliasing and reducing motion aliasing.

46. The method according to any of claims 2 and 31 and wherein said super resolution version is a slow motion version.

47. The method according to any of claims 2 and 31 and wherein said super resolution version is an up-sampled version.

48. The method according to any of claims 16 and 40 and wherein said finding comprises performing an approximate nearest neighbor search.

49. A method implementable on a computing device, the method comprising:

obtaining at least temporal super-resolution of an input video sequence using recurrence of data from said input video sequence found within an external database of examples of low-resolution / high-resolution pairs of video sequences.

50. Apparatus comprising:

a receiver to receive an input signal; and

a super resolution generator to exploit data redundancy in said input signal and to combine high frequency information from at least two different scales of said input signal to generate a super resolution version of said input signal.

51. Apparatus comprising:

a receiver to receive an input signal; and

a super resolution generator to exploit recurrence of data from said input signal in at least two different scales of at least one reference signal to extract and to combine high frequency information from a plurality of scales of said at least one reference signal to generate a super resolution version of said input signal.

52. Apparatus comprising:

a receiver to receive a single input video sequence; and

a super resolution generator to generate a super resolution version of said single input video sequence in at least the temporal dimension.

53. The apparatus according to claim 52 wherein said generator exploits recurrence of data from said single input video sequence in at least one scale of at least one of: said single input video sequence and at least one reference video sequence which is not a different simultaneous recording of the same scene.

54. A method implementable on a computing device, the method comprising:

generating a super resolution version of a single input audio signal by exploiting recurrence of data from said single input audio signal in at least one scale of at least one of: said single input audio signal and at least one reference audio signal, to recover at least high frequency information.

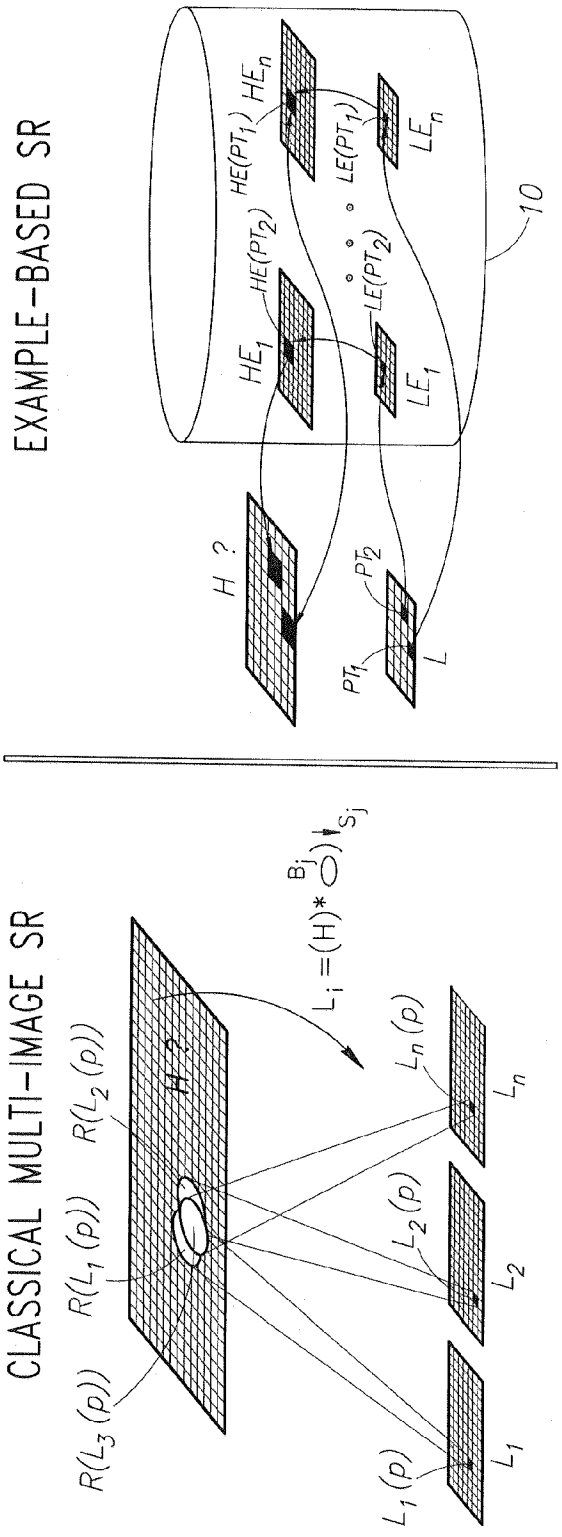

EXAMPLE-BASED SR
HE_n
HE(PT_1)
HE(PT_2)
HE_1
LE(PT_1)
LE(PT_2)
LE_n
LE_1
10
H ?
PT_2
PT_1
L
CLASSICAL MULTI-IMAGE SR
L_i =(H)* B_j ↓S_j
R(L_2 (p))
R(L_1 (p))
R(L_3 (p))
H ?
L_n (p)
L_n
L_2 (p)
L_2
L_1 (p)
L_1

FIG.1

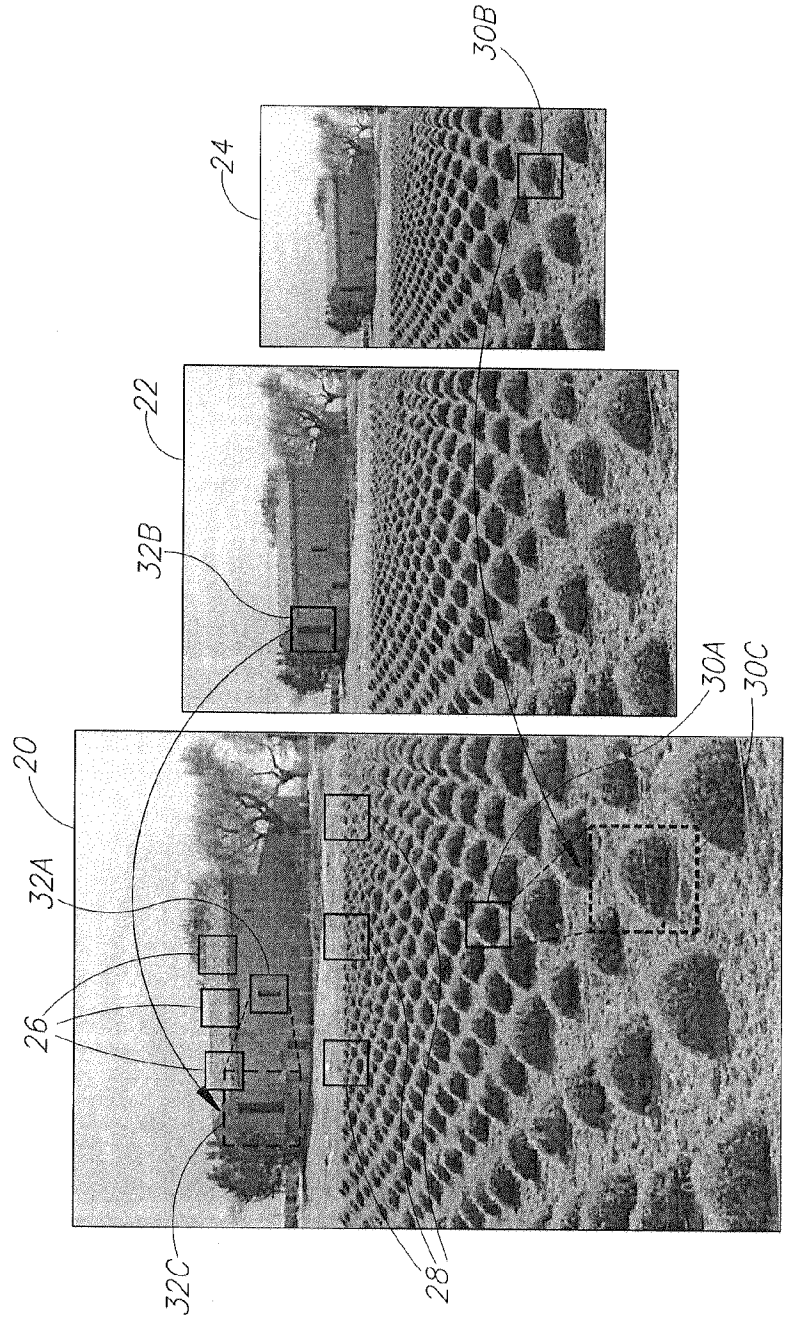


FIG.2

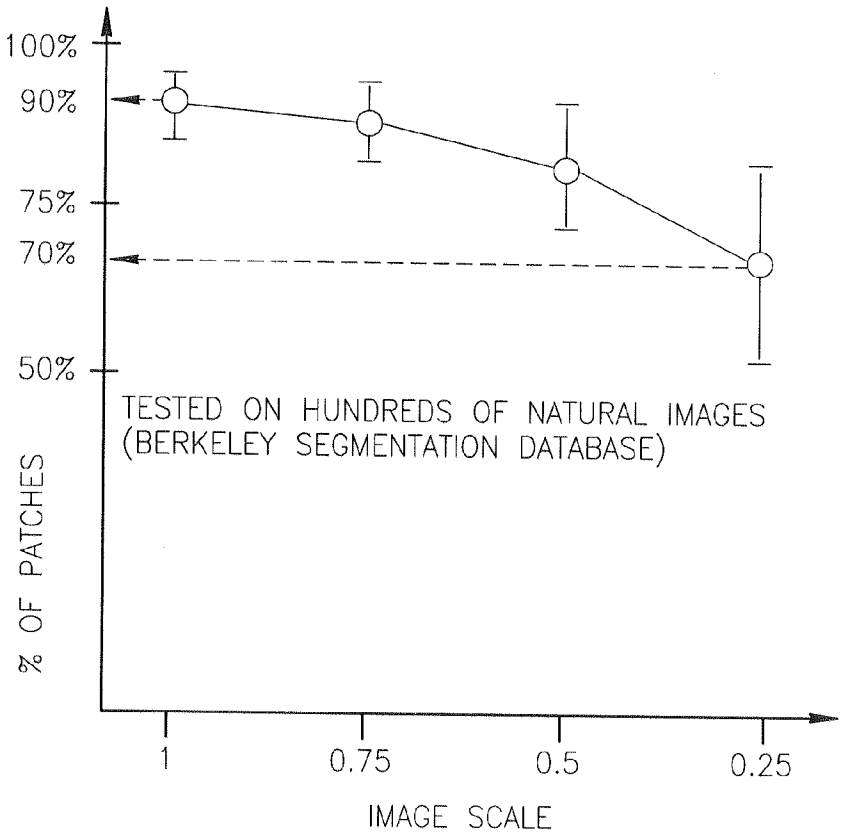

100%
90%
75%
70%
50%
% OF PATCHES
TESTED ON HUNDREDS OF NATURAL IMAGES
(BERKELEY SEGMENTATION DATABASE)
1
0.75
0.5
0.25
IMAGE SCALE

FIG.3

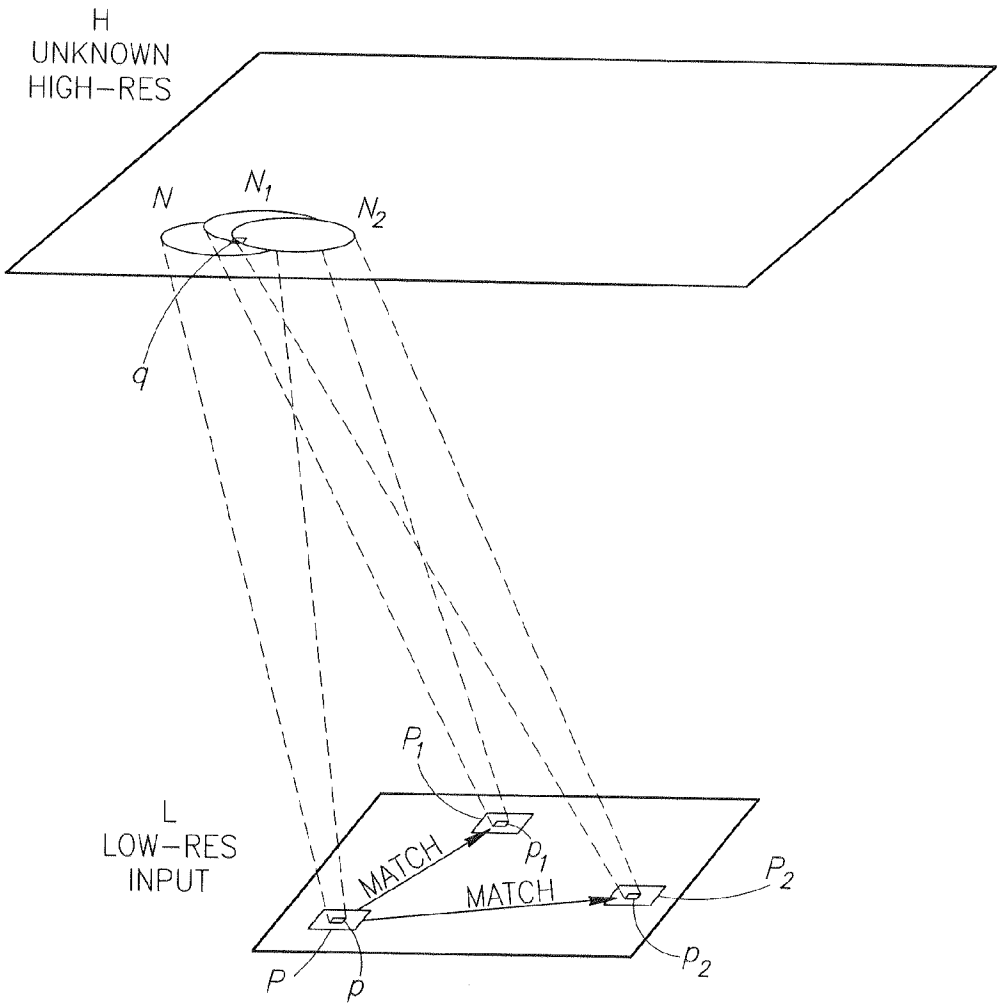

H
UNKNOWN
HIGH–RES
N
N_1
N_2
q
P_1
L
LOW–RES
INPUT
MATCH
p_1
MATCH
P_2
p_2
P
p

FIG.4A

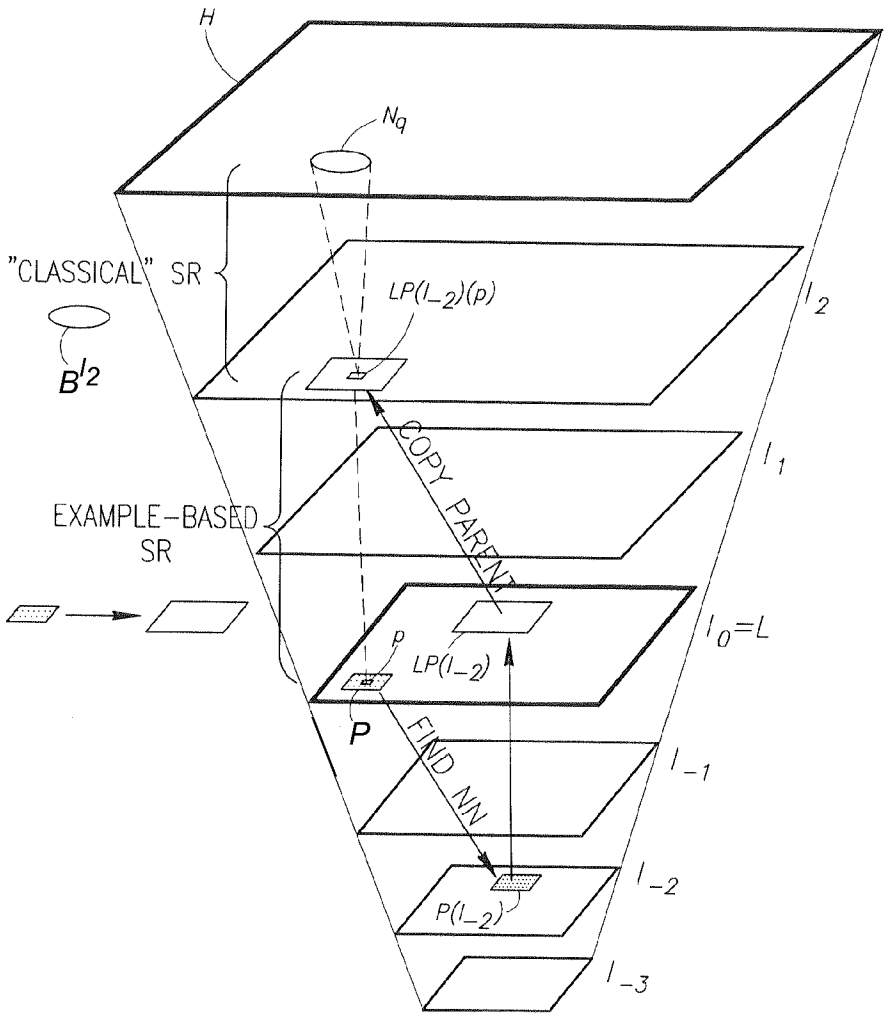

H
N_q
"CLASSICAL" SR
B^{I_2}
$LP(I_{-2})(p)$
I_2
I_1
COPY PARENT
EXAMPLE-BASED SR
$I_0=L$
p
$LP(I_{-2})$
P
FIND NN
I_{-1}
I_{-2}
$P(I_{-2})$
I_{-3}

FIG.4B

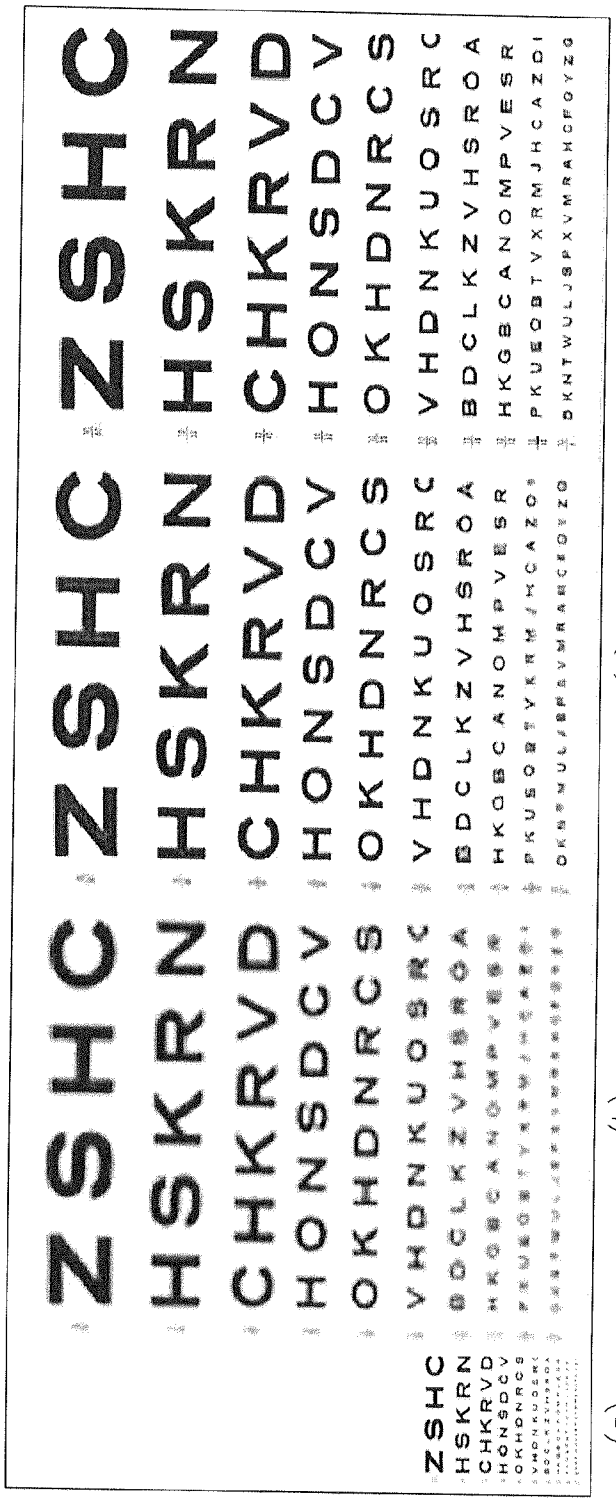

ZSHC
HSKRN
CHKRVD
HONSDCV
OKHDNRCS
VHDNKUOSRC
BDCLKZVHSROA
HKGBCANOMPVESR
PKUEOBTVXRMJHCAZDI
DKNTWULJSPXVMRAHCFOYZG
(a)
(b)
(c)
(d)

FIG.5

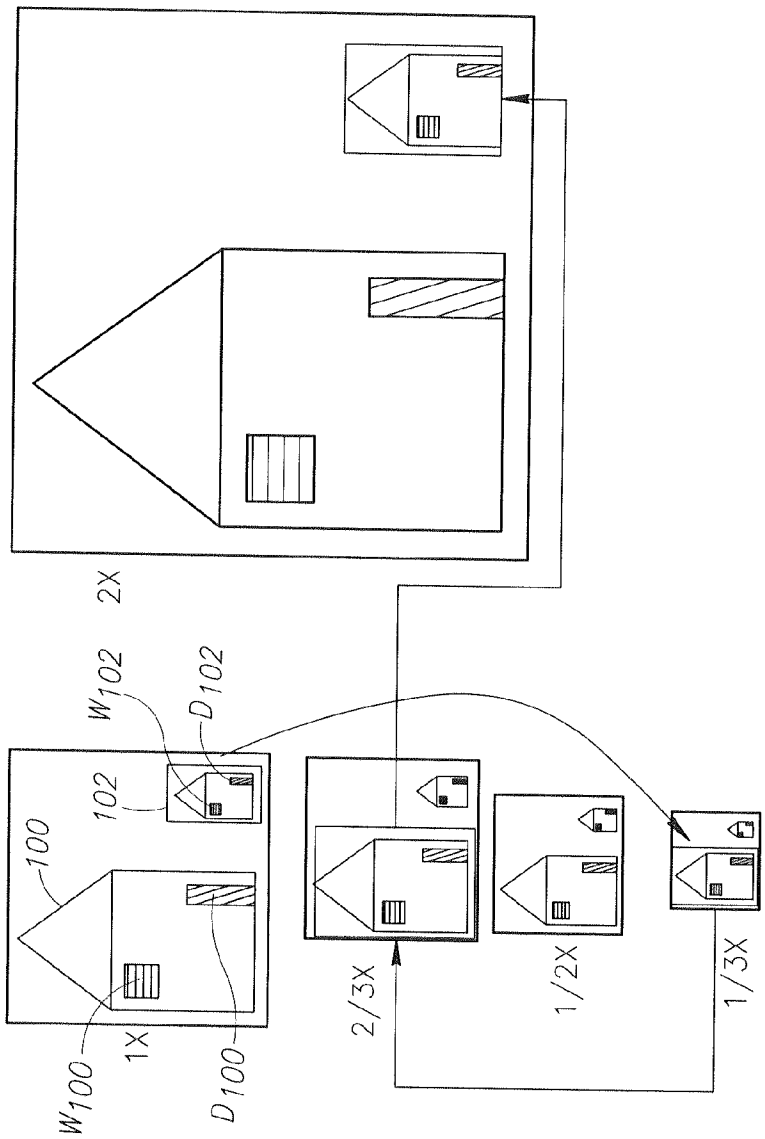


FIG.6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB 10/51738

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - G06K 9/32 (2010.01)
USPC - 382/299
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
USPC: 382/299

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC: 358/1.9; 348/14.01,14.12; 382/299,298,284,276; 386/117,107 (text search)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubWest (PGPB, USPT, EPAB, JPAB), Google,
Search terms used: image, exploiting, redundancy, recurrence, single, superresolution, super, resolution, data, redundan, combin, join, high, frequenc, input, recur, dyadic, gap, scal, temporal, reference, frequenc, transform, rotat, reflect, downscal, upscal, affine, resolut

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US 2005/0057687 A1 (Irani et al.) 17 March 2005 (17.03.2005), para [0009]-[0016], [0023]-[0030], [0042], [0048]-[0069], [0078]-[0090], [0099]-[0104], [0112]	1-12, 16-54 ------- 13-15
Y	US 2009/0046952 A1 (Ben-Ezra et al.) 19 February 2009 (19.02.2009), para [0006], [0020], [0023], [0026], [0053], [0061]	13-15
A	US 2008/0240503 A1 (Okada et al.) 02 October 2008 (02.10.2008), entire document	1-54
A	US 2006/0013332 A1 (Rayburn) 19 January 2006 (19.01.2006), entire document	1-54
A	US 2007/0296833 A1 (Corcoran et al.) 27 December 2007 (27.12.2007), entire document	1-54

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
18 September 2010 (18.09.2010)	24 SEP 2010

Name and mailing address of the ISA/US	Authorized officer:
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774