Image warping techniques

Geometric transformations between the source image and the target image

Mihalas E Mrs (PG/T - Elec Electronic Eng) | Computer Vision | November 3, 2015

Table of Contents

Abstract

This paper addresses the subject of *digital image warping* and covers the basic concepts of image warping techniques such as geometric transformations, digital image mapping, image reconstruction and antialiasing.

Warping digital images involves the change of the domain of the image through a mathematical function which establishes a relationship between the source image and the target image. It is possible to compute the values of the pixels of the target image by implementing either the forward or the backward mapping. As pixels can assume only integer values both mapping techniques require an interpolation process in order to reconstruct the image, which assumes fractional values as a result of various geometric transformations. In this paper the interpolation methods that are discussed and used are the neighbor interpolation, the bilinear interpolation and the Gaussian interpolation.

The performance of the image warp includes also a resampling process which may introduce aliasing. Therefore, it is necessary to apply a low-pass filter, which band limits the source image prior to point sampling it onto the target image, in order to prevent aliasing artifacts. Furthermore it is possible to leave source image information outside the boundaries of the target image when applying rotation, translation or up-scaling transformations. In this case it is necessary to redefine the borders of the target image by doing the out of bound correction.

Even if every shrewdness is adopted in order to create an output image identical to the input image, they will never be the same. However it is possible to compute the degree of variance between the two images by implementing an image quality measure. The measured error should be very small in order to visualize an output image fairly similar to the input image.

Introduction

A digital image is a matrix of integer numbers, called pixels, obtained by sampling the 2D space on a regular grid, named raster, and by quantizing each sample so that it represents the brightness of a given colour. The sampling process results to be very critical in most of the cases because it can introduce aliasing. Accordingly, it is necessary to apply a low-pass filter to the input image prior to any image mapping in order to reflect accurately the information content, which is mapped onto the output image.

The warping technique requires two domains, the source image and the target image, and a mathematical function, which links them together. Two strategies can be adopted in order to map the quantized values of the source image onto the target image, these are forward mapping and backward mapping. The former method is quite naïve and inaccurate and it consists of copying each pixel (x, y) of the source image onto the pixel (x', y') of the target image. However, this method raises two type of problems: holes and overlaps. On the contrary, the latter method consists of projecting each pixel (x', y') of the target image into the source image and transferring, under inverse transform, the content of the pixel (x, y) to the pixel (x', y'). Indeed the backward mapping gives better results than the forward mapping because, by visiting each pixel of the target image, it does not creates holes in the output image.

When an image is warped using a rigid body transformation all pixels are mapped under the same function. In order to transform an image it is necessary to multiply each coordinate (x, y, 1) by a 3x3 translation, scaling or rotation matrix. A composite transformation matrix is called an *affine transformation.* It implies that the homogeneous component of each coordinate (the third element) remains equal to one and that the points throughout the image are equidistant. It may happen that, by applying an affine transformation to the input image, portions of the image are lost because out of the bounds of the output image. Thus in order to preserve this information content it has to be applied an Out of Bounds correction.

The output image will never have a good visual quality if we do not include an interpolation step in the implementation process. Among different interpolation techniques, this paper discusses the nearest neighbour, the bilinear and the Gaussian interpolations.

The nearest neighbour interpolation is the simplest interpolation from the computational point of view. It assigns to each interpolated output pixel the value of the nearest sample point in the input image. Even if this technique has a very low computational cost, it generates an image with a poor visual quality because sampling artefacts are introduced in the output image. The bilinear interpolation, instead, computes the colour of non-integer coordinates by considering the linear combination of the four "closest" raster values. It assumes the image being a surface and, by approximating the surface as piecewise planar, performs the interpolation within an arbitrary quadrilateral. Consequently, it offers a better quality than the nearest neighbour interpolation. However, the Gaussian interpolation is the highest quality interpolation because it considers all pixels inside the raster. In fact the Gaussian distribution has a bell curve shape which means that its influence decades considerably with the distance from the point to be interpolated. Therefore, the interpolated colour of a certain point derives from the weighted sum of all pixels, with very little contribution coming from the distant pixels.

The interpolation technique is a powerful tool that creates a uniform and smoothed image, despite that it never creates a target image identical to the source image. However, it is possible to compute how different the two images are by measuring the root mean squared error.

Forward and backward mapping

Forward mapping uses the mapping function M to copy each discrete pixel of the input image onto the real output image positions, which may become a fraction after the computation of the affine transformation. Therefore, this non-unique match of the pixels from the two images raises several problems and affects the visual quality of the output image.

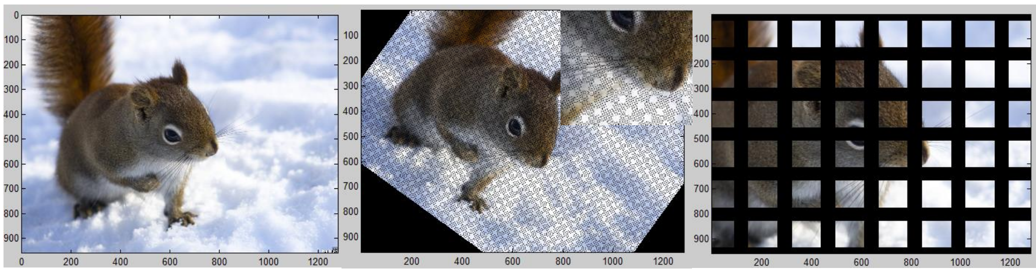

Figure 1. [From left to right] The original 960x1280 input image. The image rotated by π/5 degrees. The image scaled up by factor 1.5.

The Backward mapping, instead, uses the inverse mapping function M to project the coordinates from the target image into the source image. The affine transformation generates fractional values, alike in the forward mapping process. Therefore, it is necessary to include an interpolation stage in order to restore the discrete values of the source image positions. Once matched the positions of the two images, the content information from the source image is passed to the target image.

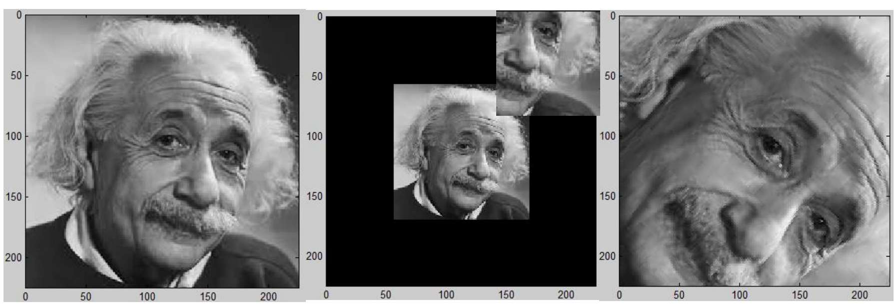

Figure 2. [From left to right] The original 225x225 input image. The image scaled down by factor 0.5. The image rotated by π/4 degrees and scaled up by factor 1.5.

Bilinear and Gaussian interpolation

The interpolation techniques are the fundamental tools in image processing because they can improve considerably the visual quality of the output image. It makes sense to talk about bilinear and Gaussian interpolations only if they are used in backward mapping because the pixel value of the target image is extracted from the pixels that surround the point of interest in the source image.

The bilinear interpolation assumes that the source image has piecewise planar patches and implements the four-corner mapping problem for quadrilaterals. Namely, it computes the intermediate values of the two horizontal sides of the quadrilateral and then the intermediate value of the line passing through these two points. The last value defines the fractional point colour that gathers the colour information of the four neighbouring pixels.

Figure 3. [From left to right] The original 201x251 input image. The bilinear interpolated image scaled up by factor 2 and rotated by π/4 degrees. The original 225x225 input image. The bilinear interpolated image rotated anticlockwise by π/4 degrees.

The Gaussian interpolation, instead, is a linear combination of Gaussian bell-shaped 2D curve with the input image pixels. First, it weights the input matrix values based on Gaussian probability distribution. Then it sums the values along all the columns and all the rows of the matrix and finally it divides the sum by the total number of matrix elements and assigns the obtained value to the output pixel. The Gaussian is centred on the point to be interpolated and the spread of the Gaussian bell curve is adjusted in order to supervise the number of influent values for the colour interpolation and the smoothness level of the output image.

Figure 4. [From left to right] The original 225x225 input image. The Gaussian interpolated image rotated clockwise by π/4 degrees with the sigma parameter equal to 5 in the first case and equal to 10 in the last case.

Pre-filtering effect

The warping process consists of re-sampling the image and it could be affected by aliasing if the frequencies in the signal are not sampled correctly. In fact, aliasing is due to under sampling, namely sampling below the Nyquist frequency, which causes high frequency components in the signal being undistinguishable from the lower frequency components. Therefore, aliasing occurs when an image is confined and has to adapt itself to an inadequate resolution, for example, it happens when the image is downscaled or rotated.

Aliasing artifacts appears in a high frequency signal as a low frequency waves which can not be eliminated from the image after the image mapping. Therefore, it is necessary to apply a low pass filter to the input image prior to any warping stage requiring scale reduction or rotation in order to reduce the value of the highest frequency in the signal and respectively to reduce the Nyquist frequency. After the anti-aliasing implementation, the image appears a little blurred but it maintains intact its frequency content.

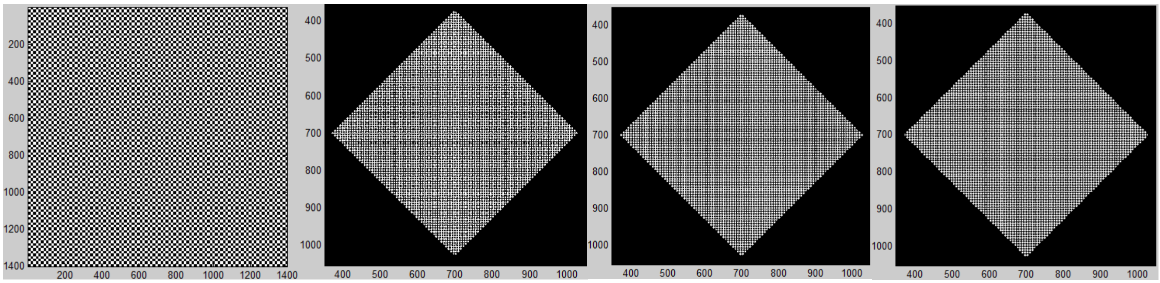

Figure 4. [From left to right] Example of a 1400x1400 chequerboard with squares of size 14 pixels. The same image scaled down by factor 1/3 and rotated π/4 degrees. The same image but pre-filtered with a box filter (3rd image) and a Gaussian filter (4th image).

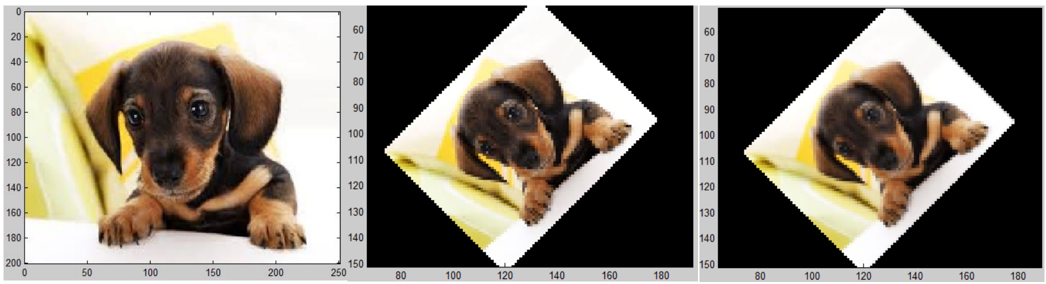

Figure 5. [From left to right] The 201x251 input image. The image bilinear interpolated, scaled down by factor 1/3 and rotated π/4 degrees. The image with the pre-filtering step implemented.

Out of Bounds Correction

When warping an image it usually happens that, the output image inherits the input image size regardless of the new boundaries determined by the affine transformation. In order to preserve the image parts that are in the negative domain or beyond the established boundaries, it is necessary to apply an out of bounds correction. It implies the calculation of the new coordinates of the warped image corners and the shifting, with a translation transformation, of all the pixels from the negative domain to the positive domain. Then the target image must be enlarged in order to include all four corners of the warped image. After the bounds correction the output image contains all the information present in the source image.

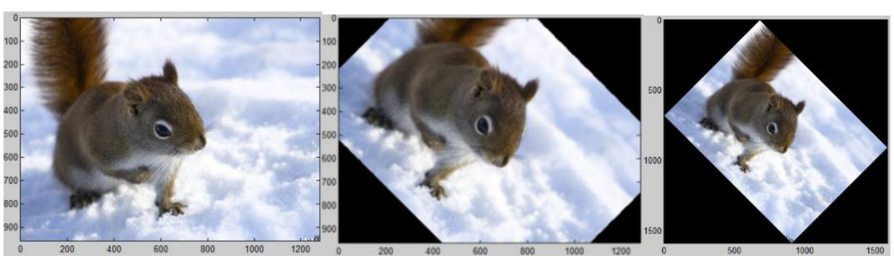

Figure 6. [From left to right] The 960x1280 input image. Image without the bounds correction. The same image after the out of bounds correction.

Measurement of error

The different interpolation techniques create target images, which are slightly different as regards the visual quality. It is possible to use an image quality metric only when the two images are of the same size and have the same content information. Therefore, it is necessary to invert the transformation process for the warped image in order to recover the “original” image. Only at this point, it is possible to compute the difference between the original image and the warped image by using the mean squared error (MSE) formula.

Figure 7. [From left to right] Original image. Nearest neighbour interpolated image. Bilinear interpolated image. Gaussian interpolated image.

The Mean Square Error between the original image and each of the interpolated image is:

1. MSE= 4.2% between the original image and the nearest neighbour interpolated image;
2. MSE=1.83% between the original image and the bilinear interpolated image;
3. MSE=20.15% between the original image and the Gaussian interpolated image.

Conclusion

The forward mapping is a simple and inaccurate technique. It has several problems as the presence of holes, namely the non-visited pixels in the output image, the presence of fractional numbers, which must be interpolated to the nearest neighbour in order to obtain a discrete value, and the loss of content at the four corners of the input image because they land outside the boundaries of the output image. On the contrary, the backward mapping visits every pixel of the target image and the holes now occur in the source image because some values from the input image are discarded while computing the values of the target image. Moreover, the interpolation techniques as bilinear and Gaussian can be implemented only through the backward mapping.

The bilinear interpolation is a valid interpolation technique, which creates target images with a low MSE. However, it adopts piecewise planar approximations which does not correspond to real variation of the colour intensity. The Gaussian interpolation, instead, gathers the weighted information from the entire matrix and assigns to the output image pixels an accurate colour value. Anyway, the weighting process affects a lot the high frequency content of the target image and it can be seen in the examples above that the images are too smoothed even when a small sigma is selected.

The pre-filtering stage is essential. In fact, the image changes its content when the high frequencies overlap with the low frequencies and this causes an irreparable artefact. Therefore, it is important to obey the Nyquist rule, namely to re-sample the original image with at least double the sampling rate of the highest frequency.

The Out of Bounds correction allows retrieving parts of the image that would have been lost in the standard mapping process. It builds an output image which height and width values are determined by computing the location of the four corners of the transformed input image.

Many objective measures of quality require a distortion-free copy of the target image that can be used for comparison with the original image. The size of the two image matrices must be identical because the difference is computed between corresponding pixels in the two images. The Mean Squared Error is a measure of how similar two images are. Smaller is the MSE, better appears the resemblance between the images. In fact, in the example above it can be noticed that the bilinear interpolated image is the most similar to the original image and it has a low MSE.

REFERENCES

- J. Collomosse, *Computer Vision and Pattern Recognition*, Digital Image Warping Lecture Notes, Guildford: University of Surrey, 2015.
- R.A. van der Stelt, “Digital Image Warping, abstract of Digital Image Warping by George Wolberg”, IEEE Computer Society Press, October 1996.
- M.S.Nixon and A.S.Aguado, Feature Extraction & Image Processing, Newnes, 2002.