

CLAIMS

1. An image processing method, comprising:

acquiring binocular images, the binocular images comprising a first image and second image which are shot for the same object in the same scenario;

obtaining first feature maps of the binocular images, first depth maps of the binocular images and second feature maps fusing an image feature and depth feature of the binocular images;

performing feature fusion processing on the binocular images, the first feature maps of the binocular images, the first depth maps of the binocular images and the second feature maps to obtain fused feature maps of the binocular images; and

performing optimization processing on the fused feature maps of the binocular images to obtain deblurred binocular images.

2. The method of claim 1, wherein obtaining the first feature maps of the binocular images comprises:

performing first convolution processing on the first image and the second image respectively to obtain first intermediate feature maps respectively corresponding to the first image and the second image;

performing second convolution processing on the first intermediate feature maps of the first image and the second image respectively to obtain second intermediate feature maps of multiple scales respectively corresponding to the first image and the second image; and

performing residual processing on the second intermediate feature maps of each scale of the first image and the second image respectively to obtain first feature maps respectively

corresponding to the first image and the second image.

3. The method of claim 2, wherein performing first convolution processing on the first image and second image of the binocular images respectively to obtain the first intermediate feature maps respectively corresponding to the first image and the second image comprises:

performing convolution processing on the first image and the second image respectively by use of a first preset convolution kernel and a first convolution step to obtain the first intermediate feature maps respectively corresponding to the first image and the second image.

4. The method of claim 2 or 3, wherein performing second convolution processing on the first intermediate feature maps of the first image and the second image respectively to obtain the second intermediate feature maps of the multiple scales respectively corresponding to the first image and the second image comprises:

performing convolution processing on the first intermediate feature maps of the first image and the second image according to preset multiple different first atrous rates respectively to obtain second intermediate feature maps respectively corresponding to the multiple first atrous rates.

5. The method of any one of claims 2-4, wherein performing residual processing on the second intermediate feature maps of each scale of the first image and the second image respectively to obtain the first feature maps respectively corresponding to the first image and the second image comprises:

concatenating second intermediate feature maps of the multiple scales of the first image respectively to obtain a first concatenated feature map, and concatenating second intermediate feature maps of the multiple scales of the second image respectively to obtain a second concatenated feature map;

performing convolution processing on the first concatenated feature map and the second concatenated feature map respectively; and

performing addition processing on a first intermediate feature map of the first image and the first concatenated feature map subjected to convolution processing to obtain a first feature map of the first image, and performing addition processing on a first intermediate feature map of the second image and the second concatenated feature map subjected to convolution processing to obtain a first feature map of the second image.

6. The method of any one of claims 1-5, wherein obtaining the first depth maps of the binocular images and the second feature maps fusing the image feature and depth feature of the binocular images comprises:

combining the first image and the second image to form a combined view;

performing, on the combined view, third convolution processing at at least one layer to obtain a first intermediate depth feature map;

performing fourth convolution processing on the first intermediate depth feature map to obtain second intermediate depth feature maps of multiple scales; and

performing residual processing on the second intermediate depth feature and the first intermediate depth feature map to obtain first depth maps of the first image and the second image respectively, and obtaining the second feature maps according to third convolution processing at any one layer.

7. The method of claim 6, wherein performing, on the combined view, third convolution processing at the at least one layer to obtain the first intermediate depth feature map comprises:

performing, on the combined view, at least one time of convolution processing by use of a second preset convolution kernel and a second convolution step to obtain the first

intermediate depth feature map.

8. The method of claim 6 or 7, wherein performing fourth convolution processing on the first intermediate depth feature map to obtain the second intermediate depth feature maps of the multiple scales comprises:

performing convolution processing on the first intermediate depth feature map according to preset multiple different second atrous rates respectively to obtain second intermediate feature maps respectively corresponding to the multiple second atrous rates.

9. The method of any one of claims 1-8, wherein performing feature fusion processing on the binocular images, the first feature maps of the binocular images, the first depth maps of the binocular images and the second feature maps to obtain the fused feature map of the binocular images comprises:

performing calibration processing on the second image according to a first depth map of the first image in the binocular images to obtain a mask map of the first image, and performing calibration processing on the first image according to a first depth map of the second image in the binocular images to obtain a mask map of the second image;

obtaining an intermediate fused feature of each image in the binocular images based on the calibrated map and mask map corresponding to each image in the binocular images;

obtaining a depth feature fused map of each image of the binocular images according to a first depth map and second feature map of each image in the binocular images; and

correspondingly obtaining a fused feature map of each image according to a concatenation result of the first feature map of the first image, an intermediate fused feature map of the first image and the depth feature fused map of the first image in all images of the binocular images.

10. The method of claim 9, wherein performing calibration processing on the second image according to the first depth map of the first image in the binocular images to obtain the mask map of the first image and performing calibration processing on the first image according to the first depth map of the second image in the binocular images to obtain the mask map of the second image comprises:

performing warp processing on the second image by use of the first depth map of the first image in the binocular images to obtain a calibrated map of the first image, and performing warp processing on the first image by use of the first depth map of the second image to obtain a calibrated map of the second image; and

obtaining the mask maps of the first image and the second image respectively according to a difference between each image in the binocular images and a corresponding calibrated map.

11. The method of claim 9, wherein obtaining the intermediate fused feature of each image in the binocular images based on the calibrated map and mask map corresponding to each image in the binocular images comprises:

obtaining the intermediate fused feature map of the first image in a first preset manner based on a calibrated map of the first image and the mask map of the first image; and

obtaining an intermediate fused feature map of the second image in a second preset manner based on the calibrated map of the second image and the mask map of the second image.

12. The method of claim 11, wherein an expression of the first preset manner is:

$F_{views}^L = F^L \odot (1 - M^L) + W^L(F^R) \odot M^L$, where F_{views}^L represents an intermediate fused feature of the first image, $\odot$ represents multiplication of corresponding elements, $W^L(I^R)$ represents a result obtained after the second image is processed by use of the first depth map of the first image, and M^L represents the mask map of the first image; and

an expression of the second preset manner is: $F_{views}^R = F^R \odot (1 - M^R) + W^R(F^L) \odot M^R$, where F_{views}^R represents an intermediate fused feature of the second image, $\odot$ represents multiplication of corresponding elements, $W^R(F^L)$ represents a result obtained after warp processing is performed on the first image by use of the first depth map of the second image, and M^R represents the mask map of the second image.

13. The method of any one of claims 1-12, wherein performing optimization processing on the fused feature maps of the binocular images to obtain the deblurred binocular images comprises:

performing convolution processing on the fused feature maps of the binocular images to obtain the deblurred binocular images.

14. An image processing device, comprising:

an acquisition module, configured to acquire binocular images, the binocular images comprising a first image and second image which are shot for the same object in the same scenario;

a feature extraction module, configured to obtain first feature maps of the binocular images, first depth maps of the binocular images and second feature maps fusing an image feature and depth feature of the binocular images;

a feature fusion module, configured to perform feature fusion processing on the binocular images, the first feature maps of the binocular images, the first depth maps of the binocular images and the second feature maps to obtain fused feature maps of the binocular images; and

an optimization module, configured to perform optimization processing on the fused feature maps of the binocular images to obtain deblurred binocular images.

15. The device of claim 14, wherein the feature extraction module comprises an image feature extraction module, configured to perform first convolution processing on the first image and second image of the binocular images respectively to obtain first intermediate feature maps respectively corresponding to the first image and the second image,

perform second convolution processing on the first intermediate feature maps of the first image and the second image respectively to obtain second intermediate feature maps of multiple scales respectively corresponding to the first image and the second image and

perform residual processing on the second intermediate feature maps of each scale of the first image and the second image respectively to obtain first feature maps respectively corresponding to the first image and the second image.

16. The device of claim 15, wherein the image feature extraction module is further configured to perform convolution processing on the first image and the second image respectively by use of a first preset convolution kernel and a first convolution step to obtain the first intermediate feature maps respectively corresponding to the first image and the second image.

17. The device of claim 15 or 16, wherein the image feature extraction module is further configured to perform convolution processing on the first intermediate feature maps of the first image and the second image according to preset multiple different first atrous rates respectively to obtain second intermediate feature maps respectively corresponding to the multiple first atrous rates.

18. An electronic device, comprising:

a processor; and

a memory, configured to store instructions executable for the processor.

wherein the processor is configured to execute the method of any one of claims 1-13.

19. A computer-readable storage medium, in which computer program instructions are stored, the computer program instructions being executed by a processor to implement the method of any one of claims 1-13.

20. A computer program product, comprising computer program instructions, the computer program instructions being executed by a processor to implement the methods of any one of claims 1-13.