CLAIMS

1. A method for depth estimation of a monocular image, comprising:

obtaining, through a depth estimation neural network, a global feature of a monocular image according to absolute features of preset regions and relative features among the preset regions in the monocular image; and

obtaining a predicted depth map of the monocular image according to the global feature, the absolute features of preset regions and relative features among the preset regions in the monocular image,

wherein after obtaining a predicted depth map of the monocular image according to the global feature, the absolute features of preset regions and relative features among the preset regions in the monocular image, the method further comprises

performing optimization on the predicted depth map according to a longitudinal variation law of depth information of the monocular image to obtain a target depth map of the monocular image.

2. The method according to claim 1, before the obtaining a global feature of a monocular image according to absolute features of preset regions and relative features among the preset regions in the monocular image, further comprising:

performing, through a first neural network, feature extraction on the monocular image to obtain features of preset regions in the monocular image, and taking the features of the preset regions as the absolute features of the preset regions in the monocular image; and

obtaining the relative features among the preset regions in the monocular image according to the absolute features of the preset regions in the monocular image.

3. The method according to claim 2, wherein the obtaining relative features among the preset regions in the monocular image according to the absolute features of the preset regions in the monocular image comprises:

performing, through an association layer, a vector operation on the absolute features of the preset regions in the monocular image to obtain the relative features among the preset regions in the monocular image.

4. The method according to claim 2 or 3, before performing, through a first neural network, feature extraction on the monocular image, further comprising:

performing downsampling on the monocular image to obtain a monocular image having a preset dimension, wherein a dimension of the monocular image is a multiple of the preset dimension.

5. The method according to any one of claims 1 to 4, wherein the obtaining a global feature of a monocular image according to absolute features of preset regions and relative features among the preset regions in the monocular image comprises:

obtaining, through a full connection layer, the global feature of the monocular image by combining the absolute features of the preset regions and the relative features among the preset regions in the monocular image .

6. The method according to any one of claims 1 to 5, wherein the obtaining a predicted depth map of the monocular image according to the global feature and the absolute features of preset regions and relative features among the preset regions in the monocular image comprises:

performing, through a depth estimator, depth estimation according to the global feature, the absolute features of the preset regions and the relative features among the preset regions in the monocular image to obtain the predicted depth map of the monocular image.

7. The method according to claim 1, wherein the performing optimization on the predicted depth map according to a longitudinal variation law of depth information of the monocular image to obtain a target depth map of the monocular image comprises:

performing residual estimation on the predicted depth map according to the longitudinal variation law of depth information of the monocular image to obtain a residual plot of the predicted depth map; and

performing optimization on the predicted depth map according to the residual plot to obtain the target depth map of the monocular image.

8. The method according to claim 7, wherein the performing residual estimation on the predicted depth map according to a longitudinal variation law of depth information of the monocular image to obtain a residual plot of the predicted depth map comprises:

performing, through a residual estimation network, residual estimation on the predicted depth map according to the longitudinal variation law of depth information of the monocular image to obtain a residual plot of the predicted depth map; and

the performing optimization on the predicted depth map according to the residual plot to obtain a target depth map of the monocular image comprises

performing a pixel-by-pixel superposition operation on the residual plot and the predicted depth map to obtain the target depth map of the monocular image.

9. The method according to any one of claims 1 to 8, before the performing optimization on the

predicted depth map according to a longitudinal variation law of depth information of the monocular image to obtain a target depth map of the monocular image, further comprising:

obtaining the longitudinal variation law of depth information of the monocular image according to the predicted depth map.

10. The method according to claim 9, wherein the obtaining a longitudinal variation law of depth information of the monocular image according to the predicted depth map comprises:

performing, through a longitudinal pooling layer, processing on the predicted depth map to obtain the longitudinal variation law of depth information of the monocular image.

11. The method according to claim 1, wherein the performing optimization on the predicted depth map according to a longitudinal variation law of depth information of the monocular image comprises:

performing a preset number of upsamplings on the predicted depth map, obtaining the longitudinal variation law of depth information according to a predicted depth map, obtained by each upsampling, having a dimension sequentially increased by a multiple, and performing optimization on the predicted depth map, obtained by each upsampling, having a dimension sequentially increased by a multiple according to the obtained longitudinal variation law of the depth information to obtain an optimized target depth map;

wherein the optimized target depth map obtained by each of the upsamplings other than a last upsampling is taken as a predicted depth map of a next upsampling, the optimized target depth map obtained by the last upsampling is taken as the target depth map of the monocular image, and the target depth map has the same dimension as the monocular image.

12. The method according to any one of claims 1 to 11, wherein the depth estimation neural network comprises an association layer, a full connection layer, and a depth estimator, and is obtained by training the depth estimation neural network by using a sparse depth map and a dense depth map obtained by stereo matching for binocular images as marking data.

13. An apparatus for the depth estimation of a monocular image, comprising:

a depth estimation neural network, configured to obtain a global feature of a monocular image according to absolute features of preset regions and relative features among the preset regions in the monocular image, and obtain a predicted depth map of the monocular image according to the global feature and the absolute features of preset regions and relative features among the preset

regions in the monocular image; and

a second neural network, configured to after obtaining the predicted depth map of the monocular image according to the global feature, the absolute features of preset regions and relative features among the preset regions in the monocular image, perform optimization on the predicted depth map according to a longitudinal variation law of depth information of the monocular image to obtain a target depth map of the monocular image.

14. An electronic device, comprising the apparatus according to claim 13.

15. An electronic device, comprising:

a memory, configured to store executable instructions; and

a processor, configured to execute the executable instructions to implement the method according to any one of claims 1 to 12.

16. A computer program, comprising a computer-readable code that, when being executed in a device, cause a processor in the device to execute instructions for implementing the method according to any one of claims 1 to 12.

17. A computer storage medium, configured to store computer readable instructions that, when being executed, implement the method according to any one of claims 1 to 12.