

TITLE OF INVENTION: METHOD AND APPARATUS FOR IMAGE SUPER-RESOLUTION PROCESSING, AND ELECTRONIC DEVICE

Technical field

The present application relates to the field of image processing technologies, and in particular to a method and an apparatus for image super-resolution processing, and an electronic device.

Backgrounds

Image super-resolution refers to a process in which a high-resolution clear image is obtained by using a low-resolution blurred image, is a very important research problem in the field of computer vision and image processing, and is widely applied in actual scenes such as medical image analysis, biometric feature recognition, and video surveillance and security.

In recent years, benefiting from the development of the deep learning technology, the super-resolution technology based on a neural network has greatly surpassed the accuracy of the traditional super-resolution technology. From the analysis of the algorithm mechanism, the current neural network super-resolution technology uses unified super-resolution to process different regions of the image, that is, uses the same means and parameters to process different regions of the image. However, in practical applications, the different regions of most images have different degrees of blur due to the distance between the actual position of the photographed object and the focal plane of the camera. For these images, the processing results of different regions need to be balanced in the unified super-resolution processing, and consequently the overall super-resolution accuracy of the image is limited.

Summary

The embodiments of the present application provide a method and apparatus for image super-resolution processing, and an electronic device, and adaptive super-resolution processing can be performed for different regions of an image, so

that processing accuracy of image super-resolution is improved.

According to a first aspect, an embodiment of the present application provides an image super-resolution processing method applied to an electronic device, including: obtaining a to-be-processed image (corresponding to a low-resolution image I below); and performing the following steps for the to-be-processed image in a n -level super-resolution processing network: predicting a n^{th} -level blurred region in the to-be-processed image, and performing super-resolution processing on the n^{th} -level blurred region in the to-be-processed image to obtain a n^{th} -level super-resolution image; predicting n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and $(n-m)^{\text{th}}$ -level blurred regions in a $(n-m)^{\text{th}}$ -level super-resolution image, and performing super-resolution processing on the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and $(n-m)^{\text{th}}$ -level blurred regions in the $(n-m)^{\text{th}}$ -level super-resolution image to obtain the $(n-m)^{\text{th}}$ -level super-resolution image, where a value of m is 1,..., $n-2$ in turn, to obtain a second-level super-resolution image; For the second-level super-resolution image, predicting the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred regions in the second-level super-resolution image, and performing super-resolution processing on the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., first-level blurred regions in the first-level super-resolution image to obtain the first-level super-resolution image, and using the first-level super-resolution image as an output image; where the degree of blur from the n^{th} -level blurred region to the first-level blurred region is sequentially reduced, and both n and m are positive integers. The n -level super-resolution processing network is a differentiated super-resolution network 10 in the following embodiments, and that n is equal to 3 is used as an example for description below. Specifically, when n is equal to 3 and m is equal to 1, the third-level blurred region in the to-be-processed image may be a third-level blurred region in the low-resolution image I, and the third-level super-resolution image may be an image S3 below. The third-level blurred region and the second-level blurred region in the third-level super-resolution image may be a third-level blurred region and a second-level blurred region in the image S3 below, that is, second-level and third-level blurred regions (that is, 2&3-level blurred region) . The second-level super-resolution image can be an image S2 below. The third-level blurred region, the

second-level blurred region, and the first-level blurred region in the second-level super-resolution image can be a third-level blurred region, a second-level blurred region, and a first-level blurred region in the image S2 below, that is, first-level, second-level and third-level blurred regions(that is, 1&2&3-level blurred region). In addition, the first-level super-resolution image may be an image S1 below. In addition, the degree of blur from the n^{th} -level blurred region to the first-level blurred region can be represented by the depth value or the distance from the actual position in the image to the focal plane of the camera. In this way, the image depth value can be perceived, multiple regions with different degrees of blur can be distinguished for the to-be-processed image based on the depth value of the image, and super-resolution processing is performed on the regions with different degrees of blur by using different super-resolution operators. Furthermore, adaptive super-resolution processing is implemented for different regions of the image, and the processing accuracy of image super-resolution is improved.

It is understandable that "A&B" in the embodiments of the present application represents a "and" relationship between "A" and "B". For example, the above-mentioned 2&3-level blurred region represents a third-level blurred region, and a second-level blurred region.

In a possible implementation of the first aspect, a difference between the n^{th} -level blurred region in the to-be-processed image and a n^{th} -level blurred region label image makes a value of a n^{th} region loss function in a corresponding predefined numerical range; a difference between the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and $(n-m)^{\text{th}}$ -level blurred regions in the $(n-m)^{\text{th}}$ -level super-resolution image and a $(n-m)^{\text{th}}$ -level blurred region label image makes a value of a $(n-m)^{\text{th}}$ -level region loss function in a corresponding predefined numerical range; and a difference between the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred regions in the first-level super-resolution image and the first-level blurred region label image makes a value of a first region loss function in a corresponding pre-defined numerical range; where the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred region label images are obtained by dividing the preset image according to n depth values ranging from large to small, and the n^{th} -level,

$(n-1)^{\text{th}}$ -level,..., and first-level blurred region label images are all single-channel images having the same size as the to-be-processed image. For example, when n is equal to 3, the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred region label images may correspond to a third-level blurred region label 12 below, a second-level and third-level blurred region label (that is, 2&3-level blurred region label)16 and a first-level, second-level and third-level blurred region label (that is, 1&2&3-level blurred region label) 20; and the n^{th} -level...and first-level region loss functions can correspond to DeepLoss3, DeepLoss2, and DeepLoss1 below. In addition, each of the predefined numerical ranges can be determined according to actual needs, and is not specifically limited in the present application, such as a numerical range between 0 and 0.5.

In a possible implementation of the first aspect, the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred region label images are obtained by using a DenseDepth network. For example, when n is equal to 3, for the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred region label images, the three levels of blurred regions can be obtained through division by a DenseDepth model according to the distances from the positions with different depth values in the given image (that is, the above to-be-processed image) to the focal plane, that is, a third-level blurred region label 12, a second-level and third-level blurred region label 16, and a first-level, second-level and third-level blurred region label 20.

In a possible implementation of the first aspect, a difference between the n^{th} -level super-resolution image and a n^{th} -level blurred super-resolution label image makes a value of a n^{th} super-resolution loss function in a corresponding predefined numerical range; a difference between the $(n-m)^{\text{th}}$ -level super-resolution image and a $(n-m)^{\text{th}}$ -level blurred super-resolution label image makes a value of a $(n-m)^{\text{th}}$ super-resolution loss function in a corresponding predefined numerical range; and a difference between the first-level super-resolution image and a first-level blurred super-resolution label image makes a value of a first super-resolution loss function in a corresponding predefined numerical range; where the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred super-resolution label images are obtained based on the fusion of

samples in a preset data set, and each sample in the preset data set includes a high-resolution image and n blurred images with different depth values, and the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., first-level blurred super-resolution label images have the same size as the to-be-processed image, and are all RGB multi-channel images. When $n=3$, the n^{th} , $(n-1)^{\text{th}}$,..., and first-level blurred super-resolution label images can be a third-level blurred region super-resolution module 13, a second-level and third-level blurred region super-resolution module(that is, 2&3-level blurred region super-resolution module) 17, and a first-level, second-level and third-level blurred region super-resolution module(that is, 1&2&3-level blurred region super-resolution module) 21 below. The n^{th} ... and first super-resolution loss functions may correspond to SRLoss_3 , SRLoss_2 , and SRLoss_1 below. The preset data set may be a realSR data set.

In a possible implementation of the first aspect, before the obtaining a to-be-processed image, the method further includes: training the n -level super-resolution processing network (that is, a differentiated super-resolution network 10 below) by using the samples in the preset data set; where in the training process, an input image of the n -level super-resolution processing network is obtained by processing the samples in the preset data set based on a target weight distribution relationship; during the training process, the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred super-resolution label images are respectively obtained by processing the samples in the preset data set based on n^{th} , $(n-1)^{\text{th}}$,..., and first weight distribution relationships; and each weight distribution relationship is used for one or more images to be fused in each sample, and a weight of each image to be fused; and the n^{th} , $(n-1)^{\text{th}}$,..., and first weight distribution relationships indicate that the smaller the depth value, the greater the weight of the blurred image. It can be understood that a depth estimation image (that is, an image including the depth value) can be extracted from the above target high-resolution pictures by using the pre-trained DenseDepth model (DenseDepth); a certain depth is randomly set as the focal plane of the camera (such as a focal plane in which the second-level blurred region is located is the focal plane of the camera); weighted averaging is performed on three blurred pictures to obtain

the fused image (that is, each super-resolution label image), where the closer the region to the focal plane of the camera, the higher the image weight of the first-level blurred region during fusion; the farther the region from the focal plane of the camera, the higher the image weight of the third-level blurred region during fusion.

In a possible implementation of the first aspect, when $n=3$, for a sample of a target high-resolution image, a first blurred image, a second blurred image, and a third blurred image included in the preset data set, the target weight relationship is used to indicate that the set output image includes the following content: a first-level blurred region obtained by fusion of a weight α_1 *a target high-resolution image + a weight β_1 *the second blurred image, a second-level blurred region obtained by fusion of a weight α_2 *the first blurred image + a weight β_2 *the second blurred image; a third-level blurred region obtained by fusion of a weight α_3 *the second blurred image + a weight β_3 *the third blurred image; a third weight relationship is used to indicate that a third-level blurred super-resolution label is set to include the following content: a first-level blurred region obtained by fusion of a weight α_1 *a target high-resolution image + a weight β_1 * the second blurred image, a second-level blurred region obtained by fusion of a weight α_2 *the first blurred image + a weight β_2 * the second blurred image; a third-level blurred region obtained by fusion of a weight α_4 *the second blurred image + a weight β_4 *the second blurred image; a second weight relationship is used to indicate a second-level blurred super-resolution label is set to include the following content: a first-level blurred region obtained by fusion of a weight α_1 *a target high-resolution image + a weight β_1 *the second blurred image, a second-level blurred region and a third-level blurred region obtained by fusion of α_5 *a target high-resolution image + a weight β_5 * the first-blurred image; and a first weight relationship is used to indicate that a first-level blurred super-resolution label is set to include the following content: the whole image directly uses the target high-resolution image; where the depth values of the first blurred image, the second blurred image, and the third blurred image sequentially increase. The specific value of each of the weights can be determined according to actual needs, and is not specifically limited in the present application. It is understandable that "*" in the

embodiment of the present application represents a multiplication sign.

According to a second aspect, an embodiment of the present application provides an image super-resolution processing apparatus applied to an electronic device, including: an obtaining module, configured to obtain a to-be-processed image; and an execution module, configured to perform the following steps for the to-be-processed image in a n -level super-resolution processing network: predict a n^{th} -level blurred region in the to-be-processed image, and perform super-resolution processing on the n^{th} -level blurred region in the to-be-processed image to obtain a n^{th} -level super-resolution image; predict n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and $(n-m)^{\text{th}}$ -level blurred regions in a $(n-m)^{\text{th}}$ -level super-resolution image, and perform super-resolution processing on the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and $(n-m)^{\text{th}}$ -level blurred regions in the $(n-m)^{\text{th}}$ -level super-resolution image to obtain the $(n-m)^{\text{th}}$ -level super-resolution image, where a value of m is 1,..., $n-2$ in turn, to obtain a second-level super-resolution image; and for the second-level super-resolution image, predict the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred regions in the second-level super-resolution image, perform super-resolution processing on the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., first-level blurred regions in the first-level super-resolution image to obtain the first-level super-resolution image, and use the first-level super-resolution image as an output image; where the degree of blur from the n^{th} -level blurred region to the first-level blurred region is sequentially reduced, and both n and m are positive integers.

In a possible implementation of the second aspect, a difference between the n^{th} -level blurred region in the to-be-processed image and a n^{th} -level blurred region label image makes a value of a n^{th} region loss function in a corresponding predefined numerical range; a difference between the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and $(n-m)^{\text{th}}$ -level blurred regions in the $(n-m)^{\text{th}}$ -level super-resolution image and a $(n-m)^{\text{th}}$ -level blurred region label image makes a value of a $(n-m)^{\text{th}}$ -level region loss function in a corresponding predefined numerical range; and a difference between the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred regions in the first-level super-resolution image and the first-level blurred region label image makes a value of a first region loss

function in a corresponding pre-defined numerical range; where the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred region label images are obtained by dividing the preset image according to n depth values ranging from large to small, and the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred region label images are all single-channel images having the same size as the to-be-processed image.

In a possible implementation of the second aspect, the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred region label images are obtained by using a DenseDepth network.

In a possible implementation of the second aspect, a difference between the n^{th} -level super-resolution image and a n^{th} -level blurred super-resolution label image makes a value of a n^{th} super-resolution loss function in a corresponding predefined numerical range; a difference between the $(n-m)^{\text{th}}$ -level super-resolution image and a $(n-m)^{\text{th}}$ -level blurred super-resolution label image makes a value of a $(n-m)^{\text{th}}$ super-resolution loss function in a corresponding predefined numerical range; and a difference between the first-level super-resolution image and a first-level blurred super-resolution label image makes a value of a first super-resolution loss function in a corresponding predefined numerical range; where the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred super-resolution label images are obtained based on the fusion of samples in a preset data set, and each sample in the preset data set includes a high-resolution image and n blurred images with different depth values, and the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., first-level blurred super-resolution label images have the same size as the to-be-processed image, and are all RGB multi-channel images.

In a possible implementation of the second aspect, the apparatus further includes: a training module, configured to: before the obtaining a to-be-processed image, train the n -level super-resolution processing network by using the samples in the preset data set; where in the training process, an input image of the n -level super-resolution processing network is obtained by processing the samples in the preset data set based on a target weight distribution relationship; during the training process, the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred super-resolution label images are respectively obtained by processing the samples in the preset data set based on n^{th} , $(n-1)^{\text{th}}$,..., and

first weight distribution relationships; and each weight distribution relationship is used for one or more images to be fused in each sample, and a weight of each image to be fused; and the n^{th} , $(n-1)^{\text{th}}$, ..., and first weight distribution relationships indicate that the smaller the depth value, the greater the weight of the blurred image.

In a possible implementation of the second aspect, when $n=3$, for a sample of a target high-resolution image, a first blurred image, a second blurred image, and a third blurred image included in the preset data set, the target weight relationship is used to indicate that the set output image includes the following content: a first-level blurred region obtained by fusion of a weight α_1 *a target high-resolution image + a weight β_1 *the second blurred image, a second-level blurred region obtained by fusion of a weight α_2 *the first blurred image + a weight β_2 *the second blurred image; a third-level blurred region obtained by fusion of a weight α_3 *the second blurred image + a weight β_3 *the third blurred image; a third weight relationship is used to indicate that a third-level blurred super-resolution label is set to include the following content: a first-level blurred region obtained by fusion of a weight α_1 *a target high-resolution image + a weight β_1 * the second blurred image, a second-level blurred region obtained by fusion of a weight α_2 *the first blurred image + a weight β_2 * the second blurred image; a third-level blurred region obtained by fusion of a weight α_4 *the second blurred image + a weight β_4 *the second blurred image; a second weight relationship is used to indicate a second-level blurred super-resolution label is set to include the following content: a first-level blurred region obtained by fusion of a weight α_1 *a target high-resolution image + a weight β_1 *the second blurred image, a second-level blurred region and a third-level blurred region obtained by fusion of α_5 *a target high-resolution image + a weight β_5 * the first-blurred image; and a first weight relationship is used to indicate that a first-level blurred super-resolution label is set to include the following content: the whole image directly uses the target high-resolution image; where the depth values of the first blurred image, the second blurred image, and the third blurred image sequentially increase.

According to a third aspect, an embodiment of the present application provides a

readable medium, where an instruction is stored on the readable medium, and when the instruction is executed on an electronic device, the electronic device is caused to execute the image super-resolution processing method according to any one of the first aspect and possible implementations of the first aspect.

According to a fourth aspect, an embodiment of the present application provides an electronic device, including: a memory configured to store an instruction executed by one or more processors of the electronic device, and a processor, being one of the processors of the electronic device, and configured to perform the image super-resolution processing method according to any one of the first aspect and possible implementations of the first aspect.

Brief description of drawings

FIG. 1 is a schematic structural diagram of a network to which an image super-resolution processing method is applied according to some embodiments of the present application;

FIG. 2 is a schematic diagram of the division of various levels of blurred regions of an image according to some embodiments of the present application;

FIG. 3 is a schematic flowchart of an image super-resolution processing method according to some embodiments of the present application;

FIG. 4 is a schematic diagram of a low-resolution image according to some embodiments of the present application;

FIG. 5 is a schematic diagram of a first-level blurred region of an image according to some embodiments of the present application;

FIG. 6 is a schematic diagram of a second-level blurred region of an image according to some embodiments of the present application;

FIG. 7 is a schematic diagram of a third-level blurred region of an image according to some embodiments of the present application;

FIG. 8 is a schematic diagram of a close-distance focus fusion image according to some embodiments of the present application;

FIG. 9 is a schematic diagram of a middle-distance focus fusion image according

to some embodiments of the present application;

FIG. 10 is a schematic diagram of a long-distance focus fusion image according to some embodiments of the present application;

FIG. 11 is a schematic structural diagram of an image super-resolution processing apparatus according to some embodiments of the present application; and

FIG. 12 is a schematic structural diagram of a mobile phone according to some embodiments of the present application.

Detailed description of embodiments

Illustrative embodiments of the present application include but are not limited to a method and an apparatus for image super-resolution processing, and an electronic device.

The image super-resolution processing method of the present application can be applied to the fields of art media, medical imaging, public security case detection, film and television production, aerial images, and cultural relic protection, and is specifically applied to a scene of image super-resolution processing in which a low-resolution blurred image is converted into a high-resolution clear image. Specifically, for a to-be-processed image for which definition needs to be adjusted, this method can perceive an image depth value, distinguish multiple regions with different degrees of blur for the to-be-processed image based on the image depth value, and perform super-resolution processing on the regions with different degrees of blur by using different super-resolution operators. Specifically, strengthened adjustment is performed on a region with a higher degree of blur (such as a background of the to-be-processed image) by using the super-resolution processing by using a heavy super-resolution operator, and slight adjustment is performed on a region with a lower degree of blur (such as a foreground of the to-be-processed image) by using super-resolution processing by using a light super-resolution operator. Therefore, it is implemented that adaptive super-resolution processing is performed for the different regions of the image, and processing accuracy of the image super-resolution is improved.

It can be understood that the image depth value in the present application indicates the depth of field of the image, for example, indicates a distance between an actual position of a photographed object in the image and a focal plane of a camera. The focal plane is a collective term for a first focal plane and a second focal plane. A plane passing through a first focus point (a front focus point or an object-side focus point) and perpendicular to a main optical axis of a system (that is, a camera system) is called the first focus plane, also known as a front focus plane or an object-side focus plane. A plane passing through a second focus (a back focus or an image-side focus point) and perpendicular to the main optical axis of the system is called the second focus plane, also known as a back focus plane or an image-side focus plane. For example, the focal plane of the camera involved in the present application may be the first focal plane of the camera system. In addition, generally, the position where a photosensitive film is installed in the camera is on the focal plane.

It should be noted that, an execution body of the image super-resolution processing method provided by the embodiments of the present application may be an electronic device, or the central processing unit (Central Processing Unit, CPU) of the electronic device, or a control module or device that is in the electronic device and that performs image super-resolution processing (or referred to as an image super-resolution processing apparatus). In the following embodiments, the electronic device is used as the execution body to illustrate the image super-resolution processing method provided in the embodiments of the present application.

It can be understood that the electronic device applicable to the present application may include but is not limited to: a mobile phone, a tablet computer, a video camera, a camera, a desktop computer, a laptop computer, a handheld computer, a notebook computer, a desktop computer, an ultra-mobile personal computer (ultra-mobile personal computer, UMPC), a netbook, as well as a cellular phone, a personal digital assistant (personal digital assistant, PDA), an augmented reality (augmented reality, AR)/a virtual reality (virtual reality, VR) device, a media player, a smart TV, a smart speaker, a smart watch, etc.

In some embodiments, the electronic device involved in the present application

has a photographing function or a camera function, and is used to obtain the to-be-processed image by photographing and perform the image super-resolution processing on the to-be-processed image above. In other embodiments, the electronic device involved in the present application is used to obtain to-be-processed images from other devices with photographing or camera functions, or to obtain the to-be-processed image from the cloud, and then perform the image super-resolution processing on the to-be-processed image.

The embodiments of the present application are further described in detail below with reference to the accompanying drawings.

FIG. 1 is a block diagram of a differentiated super-resolution network 10 of an image super-resolution processing method according to an embodiment of the present application. The differentiated super-resolution network 10 may be stored and run in an electronic device that performs the image super-resolution processing method, to support the electronic device to perform the image super-resolution processing method. As shown in FIG. 1, the network 10 includes a third-level blurred region prediction module 11, a third-level blurred region label 12, a third-level blurred region super-resolution module (that is, 3-level blurred region super-resolution module) 13, a third-level blurred super-resolution label (that is, 3-level blurred super-resolution label) 14, and a second-level and third-level blurred region prediction module (that is, 2&3-level blurred region prediction module) 15, a second-level and third-level blurred region label (2&3-level blurred region label) 16, a second-level and third-level blurred region super-resolution module (that is 2&3-level blurred region super-resolution module) 17, a second-level and third-level blurred super-resolution label (that is, 2&3-level blurred super-resolution label) 18, a first-level, second-level and third-level blurred region prediction module (that is, 1&2&3-level blurred region prediction module) 19, a first-level, second-level and third-level blurred region label (that is, 1&2&3-level blurred region label) 20, a first-level, second-level and third-level blurred region super-resolution module (that is, 1&2&3-level blurred region super-resolution module) 21, and a first-level, second-level and third-level blurred super-resolution label (that is, 1&2&3-level blurred super-resolution label) 22.

It is understandable that "A&B" in the embodiments of the present application represents a "and" relationship between "A" and "B". For example, the above-mentioned 2&3-level blurred region represents a third-level blurred region, and a second-level blurred region.

It can be understood that, in the embodiment of the present application, the electronic device that performs the image super-resolution processing method may serve as a server or a terminal electronic device.

In some embodiments, the differentiated super-resolution network 10 divides a complete image into three levels of blurred regions according to the distances from different parts of the input image to a focal plane: the part near the focal plane is a first-level blurred region (that is, a 1-level blurred region); the part that is a little farther from the focal plane is a second-level blurred region (that is, a 2-level blurred region); the part that is farthest away from the focus plane is a third-level blurred region (that is, a 3-level blurred region). The recognition of these three levels of blurred regions is obtained through supervision training on a depth estimation module (that is, a blurred region prediction module) in the differentiated super-resolution network 10. The entire differentiated super-resolution network 10 is provided with three levels of blurred region super-resolution modules. The third-level blurred region needs to be processed by all three super-resolution modules, and the second-level blurred region needs to be processed by two super-resolution modules, the first-level blurred region needs to be processed by only one super-resolution module, so as to implement the different super-resolution processing of the full-image.

Blurred region label:

More specifically, the blurred region labels above, for example, the third-level blurred region label 12, the second-level and third-level blurred region label 16, and the first-level, second-level and third-level blurred region label 20, can use a trained depth estimation model (that is, the blurred region prediction module), for example, a DenseDepth model, to obtain a depth value of a given picture (that is, an image) (a gray value ranging from 0 to 255, a single-channel image, the image has the same size

an original image). Specifically, a depth value (that is, a depth value) is randomly set to be the value of a focal plane, and the DenseDepth model divides the image into three levels of blurred regions according to the distances from the parts with different depth values in the given image to the focal plane: the part near the focal plane is a first-level blurred region; the part that is a little farther away from the focal plane is a second-level blurred region; and the part that is farthest away from the focal plane is a third-level blurred region.

It can be understood that the differentiated super-resolution network 10 provided in the present application may further include more blurred region labels, that is, more levels of blurred regions can be obtained through division according to the distances from the parts with different depth values in the input image to the focal plane. This is not specifically limited in the embodiment of the present application.

Moreover, after the image is inputted (that is, the input image of the entire differentiated super-resolution network 10), the third-level blurred region prediction module 11 is used to predict a third-level blurred region in the input image. The output image is a single-channel image that has the same size as the input image, and a pixel value of the output image indicates the third-level blurred region.

After the image (that is, the image output by the third-level blurred region super-resolution module 13) is inputted, the second-level and third-level blurred region prediction module 15 is configured to predict second-level and third-level blurred regions in the input image, that is, the second-level blurred region and the third-level blurred region. The output image is a single-channel image that has the same size as the input image, and a pixel value of the output image indicates the second-level and third-level blurred regions.

After the image (that is, the image output by the second-level and third-level blurred region super-resolution module 17) is inputted, the first-level, second-level and third-level blurred region prediction module 19 is configured to predict first-level, second-level and third-level blurred regions in the input image, that is, the first-level blurred region, the second-level blurred region, and the third-level blurred region. The output image is a single-channel image that has the same size as the input image, and

a pixel value of the output image indicates the first-level, second-level and third-level blurred regions.

The degrees of blur of the third-level blurred region, the second-level blurred region and the first-level blurred region decrease in turn, that is, the depth values decrease in turn.

Specifically, each of the blurred region prediction modules above is a spatial attention structure, and is used to perform a pooling operation on channels through which the image is input, to obtain a single-channel output image that has the same size as the input image. The pixel value of the output image indicates a blurred region of the output image.

According to some embodiments of the present application, each of the blurred region labels above may be obtained by pre-training the DenseDepth model. Specifically, during the training process, the differentiated super-resolution network 10 may use a corresponding loss function to implement supervision training to obtain each blurred region label above.

For example, as shown in FIG. 1, based on a loss function DeepLoss3 between the output image of the third-level blurred region prediction module 11 and the third-level blurred region label 12, a loss function DeepLoss2 between the output image of the second-level and third-level blurred region prediction module 15 and the second-level and third-level blurred region label 16, and a loss function DeepLoss1 between the output image of the first-level, second-level and third-level blurred region prediction module 19 and the first-level, second-level and third-level blurred region label 20, the DenseDepth model can perform supervision training to obtain the third-level blurred region label 12, the second-level and third-level blurred region label 16, and the first-level, second-level and third-level blurred region label 20.

In some embodiments, a loss function (Deeploss) can be implemented by mse. Specifically, the blurred region label (a single-channel image that has the same size as an original image, and a pixel value of the image is the corresponding depth of the pixel) corresponding to the loss function can be denoted as LD, and the output image of the blurred region prediction module (that is, a depth of field image)

is denoted as PD, Deeploss=mse(PD,LD). For example, the loss function between the third-level blurred region label 12 (denoted as LD3) and the output image of the third-level blurred region prediction module 11 (denoted as PD3) is Deeploss3=mse(PD3, LD3).

MSE (Mean Square Error), as a statistical parameter, is the mean value of the sum of squares of the errors between the predicted data and the corresponding original data. The calculation formulas are as follows:

$$MSE = \frac{SSE}{n} = \frac{1}{n} \sum_{i=1}^m w_i (y_i - \hat{y}_i)^2$$

$$SSE = \sum_{i=1}^m w_i (y_i - \hat{y}_i)^2$$

In the formulas, y_i is the real data (PD), $\hat{y}_i$ is the fitted data (that is, LD), $w_i > 0$, and SSE (and a variance) is close to 0, it indicates that the model selection and fitting are better, and the data prediction is more successful, where m is the number of pixels, n is the number of samples.

Blurred super-resolution label:

The input image of the third-level blurred region super-resolution module 13 is the image output by the third-level blurred region prediction module 11, and the third-level blurred region super-resolution module 13 is used to perform super-resolution processing on the third-level blurred region of the input image. The input image of the second-level and third-level blurred region super-resolution module 17 is the image output by the second-level and third-level blurred region prediction module 15, and the second-level and third-level blurred region super-resolution module 17 is used to perform super-resolution processing on the second-level and third-level blurred regions of the input image. The input image of the first-level, second-level and third-level blurred region super-resolution module 21 is the image output by the second-level and third-level blurred region prediction module 19, and the first-level, second-level and third-level blurred region

super-resolution module 21 is used to perform super-resolution processing on the first-level, second-level and third-level blurred regions of the input image.

In some embodiments, the main framework of the third-level blurred region super-resolution module 13, the second-level and third-level blurred region super-resolution module 17, and the first-level, second-level and third-level blurred region super-resolution module 21 is an encoder-decoder structure, and mainly involves convolution operation, the size and content of the output image are consistent with that of the input image, and both the input image and the output image are RGB channel images, but the output image is a clearer image obtained by performing super-resolution processing on the input image.

In some embodiments, spatial attention refers to applying average-pooling and max-pooling operations along the channel axis, and then concatenating them to generate an efficient feature descriptor. Then, the convolutional layer is applied to generate a spatial attention map $M_s(F)$ with a size of $R \times H \times W$, and the spatial attention map encodes where to emphasize or suppress in the image.

$$\begin{aligned} M_s(F) &= \sigma(f^{7 \times 7}([AvgPool(F) ; MaxPool(F)])) \\ &= \sigma(f^{7 \times 7}([F_{avg}^S ; F_{max}^S])) \end{aligned}$$

Specifically, the spatial attention structure aggregates the channel information of a feature map (feature descriptor) by using two pooling operations to generate two 2D maps: F_{avg}^S with a size of $1 \times H \times W$, and F_{max}^S with a size of $1 \times H \times W$, where σ indicates the sigmoid function, and $f^{7 \times 7}$ indicates a convolution operation with the filter size of 7×7 .

In some embodiments, the three blurred region super-resolution modules perform super-resolution processing by using the same super-resolution operators, or the third-level blurred region super-resolution module 13, the second-level and third-level blurred region super-resolution module 17, and the first-level, second-level and third-level blurred region super-resolution module 21 perform super-resolution processing by using the super-resolution operators whose degrees of strength sequentially decrease. This is not specifically limited in the embodiment of the

present application.

In some embodiments, a third-level blurred super-resolution label 14, a second-level and third-level blurred super-resolution label 18, and a first-level, second-level and third-level level blurred super-resolution label 22 may be obtained by fusion of image data in a preset data set. For example, each sample of the preset data set includes a high-resolution picture and three levels of blurred pictures after pixel-level alignment. As an example, the preset data set is a realSR data set. For example, a set of images in the preset data set includes a high-resolution picture HR (HR); a slightly blurred picture LR2 (LR2); a more blurred picture LR3 (LR3) and a most blurred picture LR4 (LR4). In some embodiments, HR, LR2, LR3, and LR4 are all color images in RGB formats.

More specifically, as shown in FIG. 2, for a given picture, the blurred region prediction module can distinguish three levels of blurred regions in the picture by using the depth value: the part marked with the symbol A is the first-level blurred region closest to the focal plane, the part marked with the symbol B is the second-level blurred region slightly farther from the focal plane, and the part marked with the symbol C is the third-level blurred region farthest from the focal plane.

In some embodiments, in a process of training the differentiated super-resolution network 10 by using a set of images in the preset data set includes HR; LR2; LR3 and LR4, the input image of the entire network includes the following parts: The part marked with the symbol A is obtained by the fusion of $\alpha_1 \cdot \text{HR} + \beta_1 \cdot \text{LR2}$; the part marked with the symbol B is obtained by the fusion of $\alpha_2 \cdot \text{LR2} + \beta_2 \cdot \text{LR3}$; the part marked with the symbol C is obtained by the fusion of $\alpha_3 \cdot \text{LR3} + \beta_3 \cdot \text{LR4}$. It can be understood that the values of α_1 , β_1 , α_2 , β_2 , α_3 , and β_3 are not specifically limited in the embodiments of the present application, and can be set according to actual needs. For example, assuming that the focal plane of the camera is the focal plane on which the part marked with the symbol B is located, α_1 is greater than β_1 , α_2 is greater than β_2 , and α_3 is less than β_3 . It is understandable that "*" in the embodiment of the present application represents a multiplication sign.

It can be understood that, in the present application, supervision training can be

performed on the differentiated super-resolution network 10 by using the preset data set (such as the realSR data set) to obtain the third-level blurred super-resolution label 14, the second-level and third-level blurred super-resolution label 18, and the first-level, second-level and second-level blurred super-resolution label 22.

Specifically, with reference to the example in FIG. 2, the third-level blurred super-resolution label 14 in FIG. 1 includes the following parts: The part marked with the symbol A is the same as the input image, that is, obtained by fusion of $\alpha_1 \cdot \text{HR} + \beta_1 \cdot \text{LR}_2$. The part marked with the symbol B is the same as the input image, that is, obtained by the fusion of $\alpha_2 \cdot \text{LR}_2 + \beta_2 \cdot \text{LR}_3$. The part marked with the symbol C is obtained by the fusion of $\alpha_4 \cdot \text{LR}_2 + \beta_4 \cdot \text{LR}_3$. It can be understood that the values of α_4 and β_4 are not specifically limited in the embodiments of the present application, and can be set according to actual requirements, for example, α_4 is smaller than β_4 .

With reference to the example in FIG. 2, the second-level and third-level level blurred super-resolution label 18 in FIG. 1 includes the following parts: The part marked with the symbol A is the same as the input image, that is, obtained by the fusion of $\alpha_1 \cdot \text{HR} + \beta_1 \cdot \text{LR}_2$. Both the blurred super-resolution label of the part marked with the symbol B and the blurred super-resolution label of the part marked with the symbol C are obtained by the fusion of $\alpha_5 \cdot \text{HR} + \beta_5 \cdot \text{LR}_2$.

With reference to the example in FIG. 2, the first-level, second-level and third-level blurred super-resolution label 20 in FIG. 1 includes the following parts: HR is directly used for the whole picture, that is, HR is used as the first-level, second-level and third-level blurred super-resolution label 20 by the part marked with the symbol A, the part marked with the symbol B and the part marked with the symbol C.

According to some embodiments of the present application, the blurred super-resolution labels above can be obtained by the fusion of the realSR data set. Specifically, during the training process, the differentiated super-resolution network 10 may use a corresponding loss function to implement supervise learning for each blurred super-resolution label.

For example, as shown in FIG. 1, based on a loss function SRLoss_3 between the

output image of the third-level blurred super-resolution module 13 and the third-level blurred super-resolution label 14, a loss function SR_{Loss2} between the output image of the second-level and third-level blurred super-resolution module 17 and the second-level and third-level blurred super-resolution label 18, and a loss function SR_{Loss1} between the output image of the first-level, second-level and third-level blurred super-resolution module 21 and the first-level, second-level and third-level blurred super-resolution label 22, supervision training is performed to obtain the third-level blurred super-resolution label 14 the second-level and third-level blurred super-resolution label 18 and the first-level, second-level and third-level blurred super-resolution label 22.

In some embodiments, a loss function SR_{loss} can be implemented by mse. Specifically, assuming that a high-resolution label image is HR, and the super-resolution image finally output by the differentiated super-resolution network 10 is SR, the SR_{loss} (that is, SR_{loss1}) is obtained by using traditional super-resolution $SR_{loss}=mse(SR,HR)$. The SR_{loss} in the intermediate process of the differentiated super-resolution network 10 uses the same mse function, and only HR and SR are replaced with a corresponding module output image and a label image.

It can be understood that the solid arrow shown in FIG. 1 represents the convolutional network layer, and the dotted arrow represents the loss function corresponding to the supervised training label. As an example, the blurred super-resolution labels during training are obtained by fusion of the public realSR data set, and the blurred region labels are obtained by the DenseDepth network model.

In addition, it can be understood that, in other embodiments, the differentiated super-resolution network 10 provided by the present application may further include more than three levels of blurred region prediction units and blurred region super-resolution units, which are used to perform finer division on the regions with different degrees of blur in the input image, and then refine super-resolution processing of different degrees on the regions with different degrees. In other words, for the entire input image of the differentiated super-resolution network 10, regions with degrees of blur of more than three levels can be distinguished according to the

image depth value.

Specifically, a strategy in which the image fusion is performed by using the preset data set is as follows: 1) A depth estimation map (that is, an image including a depth value) is extracted from a high-resolution image by a pre-trained DenseDepth model (DenseDepth); 2) A depth is randomly set as the focal plane of the camera; 3) Weighted averaging is performed on three blurred images to obtain the fused image. The closer the region to the focal plane of the camera, the higher the weight of the image of the first-level blurred region during fusion; the further the region from the focal plane of the camera, the higher the weight of the image of the third-level blurred region during fusion. Through this strategy, the fusion of any focal plane of the image can be implemented, that is, the image super-resolution processing method of the present application can be implemented.

According to some embodiments of the present application, as shown in FIG. 3, FIG. 3 is a flow of an image super-resolution processing method based on the differentiated super-resolution network 10 shown in FIG. 1, and the flow includes step 301-step 307:

Step 301: Input a low-resolution image I into the differentiated super-resolution network 10. The low-resolution image I is a to-be-processed image for which resolution needs to be improved.

As shown in FIG. 4, FIG. 4 is a low-resolution image I, that is, an example of a to-be-processed image (an original image), and the image is a color image in a RGB format.

Step 302: The differentiated super-resolution network 10 predicts a third-level blurred region (that is, the third-level blurred region) in the low-resolution image I by using a third-level blurred region prediction module 11 (that is, an attention module).

In a possible implementation, the differentiated super-resolution network 10 in the step 302 predicts the third-level blurred region in the low-resolution image I based on a third-level blurred region label 12 and a loss function DeepLoss3 by using the third-level blurred region prediction module 11.

According to some embodiments of the present application, when DeepLoss3 is in a corresponding predetermined value range, such as a value range of 0-0.5, it is considered that the third-level blurred region prediction module 11 accurately predicts the third-level blurred region in the low-resolution image I.

Certainly, the predetermined numerical range of each loss function listed in the embodiment of the present application is only an example, and may further be any achievable numerical range. This is not specifically limited in the embodiment of the present application.

Exemplarily, with reference to the low-resolution image I shown in FIG. 4, as shown in FIGS. 5-7, FIGS. 5-7 are respectively schematic diagrams of the first-level blurred region, the second-level blurred region, and the third-level blurred region in the low-resolution image I predicted by the differentiated super-resolution network 10.

Step 303: The differentiated super-resolution network 10 performs super-resolution processing on the third-level blurred region in the image I by using a convolution operation by using the third-level blurred region super-resolution module 13 to obtain an image S3.

Specifically, based on the third-level blurred super-resolution label 14 and the loss function SRLoss3, the differentiated super-resolution network 10 in the step 303 may perform super-resolution processing on the third-level blurred region in the image I by using a convolution operation by the third-level blurred region super-resolution module 13 to obtain the image S3.

According to some embodiments of the present application, when SRLoss3 is in a corresponding predetermined value range, such as a value range of 0-0.5, it is considered that the third-level blurred region super-resolution module 13 successfully obtains the image S3 obtained after super-resolution processing.

Step 304: The differentiated super-resolution network 10 predicts a second-level blurred region and a third-level blurred region in the image S3 by using the second-level and third-level blurred region prediction module 15 (that is, an attention module). Where, "&" means "and".

It can be understood that the second-level blurred region and the third-level blurred region in the image S3 can also be understood as the second-level blurred region and the third-level blurred region in the low-resolution image I.

In a possible implementation, based on the second-level and third-level blurred region label 16 and the loss function DeepLoss2, the differentiated super-resolution network 10 in step 304 may predict the second-level blurred region and the third-level blurred region in the image S3 by using the second-level and third-level blurred region prediction module 15.

According to some embodiments of the present application, when DeepLoss2 is in a corresponding predetermined value range, such as a value range of 0-0.5, it is considered that the second-level and third-level blurred region prediction module 15 accurately predicts the second-level blurred region and the third-level blurred region in the image S3.

Step 305: The differentiated super-resolution network 10 performs super-resolution processing on the second-level blurred region and the third-level blurred region in the image S3 by using a convolution operation by using the second-level and third-level blurred region super-resolution module 17 to obtain an image S2.

Specifically, based on the second-level and third-level blurred super-resolution label 18 and the loss function SRLoss2, the differentiated super-resolution network 10 in the step 305 may perform super-resolution processing on the second-level blurred region and the third-level blurred region in the image S3 by using a convolution operation by using the second-level and third-level blurred region super-resolution module 17 to obtain the image S2.

According to some embodiments of the present application, when SRLoss2 is in a corresponding predetermined value range, such as a value range of 0-0.5, it is considered that the second-level and third-level blurred region super-resolution module 17 successfully obtains the image S2 obtained after super-resolution processing.

Step 306: The differentiated super-resolution network 10 predicts the first-level

blurred region, the second-level blurred region, and the third-level blurred region (first-level, second-level and third-level blurred regions) in the image S2 by using the first-level, second-level and third-level blurred region prediction module 19 (that is, an attention module).

In a possible implementation, based on the first-level, second-level and third-level blurred region label 20 and the loss function DeepLoss1, the differentiated super-resolution network 10 in the step 306 predicts the first-level, second-level and third-level blurred regions in the image S2 by using the first-level, second-level and third-level blurred region prediction module 19.

According to some embodiments of the present application, when DeepLoss1 is in a corresponding predetermined value range, such as a value range of 0-0.5, it is considered that the first-level, second-level and third-level blurred region prediction module 19 accurately predicts the first-level, second-level and third-level blurred regions in image S2.

It can be understood that the first-level blurred region, the second-level blurred region, and the third-level blurred region in the image S2 can also be understood as the first-level blurred region, the second-level blurred region, and the third-level blurred region in the low-resolution image I.

Step 307: The differentiated super-resolution network 10 performs super-resolution processing on the first-level blurred region, the second-level blurred region, and the third-level blurred region in the image S2 by the first-level, second-level and third-level blurred region super-resolution module 21 by using a convolution operation to obtain an image S1, that is, the output image of the entire differentiated super-resolution network 10.

Specifically, based on the first-level, second-level and third-level blurred super-resolution label 22 and the loss function SRLoss1, the differentiated super-resolution network 10 in the step 307 may perform super-resolution processing on the first-level blurred region, the second-level blurred region and the third-level blurred region in the image S2 by using a convolution operation by the first-level, second-level and third-level blurred region super-resolution module 21 to obtain the

image S1.

According to some embodiments of the present application, when SR_{Loss1} is in a corresponding predetermined value range, such as a value range of 0-0.5, it is considered that the first-level, second-level and third-level blurred region super-resolution module 21 successfully obtains the image S1 obtained after super-resolution processing.

It can be understood that when the focal plane of the camera is randomly set to be a focal plane at a close distance, that is, the camera is focused at a close distance, the differentiated super-resolution network 10 adopts the steps 301 to 307, that is, adopts the image super-resolution processing method of the application to process the low-resolution image I shown in FIG. 4 to obtain a super-resolution image as shown in FIG. 8, that is, a close-distance focus fusion image.

When the focal plane of the camera is randomly set to be the focal plane at a middle distance, that is, when the camera is focused at the middle distance, the differentiated super-resolution network 10 uses the steps 301 to 307 to process the low-resolution image I as shown in FIG. 4 to obtain a super-resolution image shown in FIG. 9, that is, a middle-distance focus fusion image.

When the focal plane of the camera is randomly set to be the focal plane at a long distance, that is, the camera is focused at a long distance, the differentiated super-resolution network 10 uses the steps 301 to 307 to process the low-resolution image I shown in FIG. 4 to obtain a super-resolution image shown in FIG. 10, that is, a long-distance focus fusion image. The images shown in FIGS. 8-10 are all color images in RGB formats.

It can be understood that the judgment of the distance at which the focal plane of the camera is set involved in the present application can be determined based on the depth value of the image, and the settings of the close distance, the middle distance, and the long distance can be determined according to actual needs. This is not specifically limited in the embodiments of the present application.

In this way, the present application can implement the fusion of any focal plane of the image, thereby implementing high-precision image super-resolution processing.

The embodiments of the present application may divide the image super-resolution processing apparatus into one or more modules based on the method examples. For example, the modules may be divided corresponding to functions, or two or more functions may be integrated into one processing module. The integrated modules above can be implemented in the form of hardware or software modules. It should be noted that the division of modules in the embodiments of the present application is illustrative, and is only a logical function division, and there may be other division methods in actual implementation.

FIG. 11 shows a schematic diagram of a possible structure of the image super-resolution processing apparatus provided in the foregoing embodiment. The image super-resolution processing apparatus may be an electronic device such as a mobile phone or a camera with image capturing and processing functions, or a CPU of the electronic device, or a control module for performing image super-resolution processing in the electronic device.

As shown in FIG. 11, the image super-resolution processing apparatus 1100 includes: an obtaining module 1101, configured to obtain a to-be-processed image; an execution module 1102, configured to perform the following steps for the to-be-processed image obtained by the obtaining module 1101 in a n -level super-resolution processing network: predict a n^{th} -level blurred region in the to-be-processed image, and perform super-resolution processing on the n^{th} -level blurred region in the to-be-processed image to obtain a n^{th} -level super-resolution image; predict n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and $(n-m)^{\text{th}}$ -level blurred regions in a $(n-m)^{\text{th}}$ -level super-resolution image, and perform super-resolution processing on the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and $(n-m)^{\text{th}}$ -level blurred regions in the $(n-m)^{\text{th}}$ -level super-resolution image to obtain the $(n-m)^{\text{th}}$ -level super-resolution image, where a value of m is 1,..., $n-2$ in turn, to obtain a second-level super-resolution image; for the second-level super-resolution image, predict the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred regions in the second-level super-resolution image, perform super-resolution processing on the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., first-level blurred regions in the first-level super-resolution image to obtain the first-level super-resolution

image, and use the first-level super-resolution image as an output image; where the degree of blur from the n^{th} -level blurred region to the first-level blurred region is reduced in turn, and both n and m are positive integers.

It can be understood that the image-based super-resolution processing apparatus 1101 shown in FIG. 11 corresponds to the method provided in the present application, and the technical details in the specific description of the method provided in the present application are still applicable to the image super-resolution processing apparatus 1100 shown in FIG. 11. For specific description, please refer to the above contents. Details are not described herein.

According to an embodiment of the present application, an electronic device is provided. The electronic device may be a mobile phone, a tablet, a wearable device, etc., with shooting function and image processing function. The following description takes a mobile phone as an example of an electronic device.

As shown in FIG. 12, the mobile phone 10 may include a processor 110, a power module 140, a memory 180, a mobile communication module 130, a wireless communication module 120, a sensor module 190, an audio module 150, a camera 170, an interface module 160, buttons 101, and a display screen 102 and so on.

It is understandable that the structure illustrated in the embodiment of the present invention does not constitute a specific limitation on the mobile phone 10. In other embodiments of the present application, the mobile phone 10 may include more or fewer components than those shown in the figure, or combine certain components, or split certain components, or arrange different components. The illustrated components can be implemented in hardware, software, or a combination of software and hardware.

The processor 110 may include one or more processing units, for example, may include a central processing unit (CPU), an image processor GPU (graphics processing unit), a digital signal processor DSP, and a microprocessor MCU (Micro-programmed Control Unit), AI (Artificial Intelligence, artificial intelligence) processor or programmable logic device FPGA (Field Programmable Gate Array) and other processing modules or processing circuits. A storage unit may be provided in the

processor 110 to store instructions and data. In some embodiments, the storage unit in the processor 110 is the cache memory 180 (and the memory 180). For example, the processor 110 is configured to execute the image super-resolution processing method in the foregoing embodiment, and the memory 180 stores instructions corresponding to the image super-resolution processing method in the foregoing embodiment.

The power supply module 140 may include a power supply, a power management component, and the like. The power source can be a battery. The power management component is used to manage the charging of the power supply and the power supply to other modules.

The mobile communication module 130 may include, but is not limited to, an antenna, a power amplifier, a filter, an LNA (Low noise amplify, low noise amplifier), and the like. The mobile communication module 130 can provide a wireless communication solution including 2G/3G/4G/5G and the like applied on the mobile phone 10. The mobile communication module 130 may receive electromagnetic waves by an antenna, and perform processing such as filtering and amplifying the received electromagnetic waves, and transmit them to the modem processor for demodulation.

The wireless communication module 120 may include an antenna, and transmit and receive electromagnetic waves via the antenna. The wireless communication module 120 can provide applications on the mobile phone 10 including wireless local area networks (WLAN) (such as wireless fidelity (Wi-Fi) networks), Bluetooth (bluetooth, BT), and global navigation satellite systems (Global navigation satellite system, GNSS), frequency modulation (frequency modulation, FM), near field communication technology (near field communication, NFC), infrared technology (infrared, IR) and other wireless communication solutions. The mobile phone 10 can communicate with the network and other devices through wireless communication technology.

In some embodiments, the mobile communication module 130 and the wireless communication module 120 of the mobile phone 10 may also be located in the same module.

The display screen 102 is used to display human-computer interaction interfaces, images, videos, and the like. The display screen 102 includes a display panel.

The audio module 150 is used to convert digital audio information into an analog audio signal for output, or convert an analog audio input into a digital audio signal. The audio module 150 can also be used to encode and decode audio signals.

The camera 170 is used to capture still images or videos. The object generates an optical image through the lens and projects it to the photosensitive element. The photosensitive element converts the light signal into an electrical signal, and then transfers the electrical signal to ISP (Image Signal Processing, image signal processing) to convert into a digital image signal. For example, it is used to support the mobile phone 10 to collect the aforementioned image to be processed.

The interface module 160 includes an external memory interface, a universal serial bus (USB) interface, and a subscriber identification module (SIM) card interface. The external memory interface can be used to connect an external memory card, such as a Micro SD card, to expand the storage capacity of the mobile phone 10. The external memory card communicates with the processor 110 through an external memory interface to realize the data storage function.

In some embodiments, the mobile phone 10 further includes a button 101, a motor, an indicator, and the like. Among them, the button 101 may include a volume button, an on/off button, and the like.

All the embodiments of the mechanism disclosed in the present application may be implemented in hardware, software, firmware, or a combination of these implementation methods. The embodiments of the present application may be implemented as computer programs or program codes executed on a programmable system, and the programmable system includes at least one processor, a storage system (including a volatile and non-volatile memory and/or a storage element), at least one input device, and at least one output device.

The program codes can be applied to input an instruction to perform the functions described in the present application and generate output information. The output information can be applied to the one or more output devices in a known

manner. For the objective of the present application, a processing system includes any system having a processor such as a digital signal processor (DSP), a microcontroller, an application-specific integrated circuit (ASIC), or a microprocessor, etc.

The program codes can be implemented in a high-level programming language or an object-oriented programming language in order to communicate with the processing system. When needed, the program codes can also be implemented in an assembly language or a machine language. In fact, the mechanism described in the present application is not limited to the scope of any particular programming language. In either case, the language can be a compiled language or an interpreted language.

In some cases, the disclosed embodiments may be implemented in hardware, firmware, software, or any combination thereof. The disclosed embodiments may further be implemented as instructions carried on or stored by one or more transitory or non-transitory machine-readable (such as computer readable) storage media, and the instructions may be read or executed by one or more processors. For example, the instructions may be distributed by using a network or by using another computer readable medium. Therefore, the machine-readable medium may include any mechanism for storing or transmitting information in a machine (such as a computer) readable form, including but not limited to, a floppy disk, an optical disk, an optical disc, a read-only memory (CD-ROM), a magneto-optical disk, a read-only memory (ROM), a random access memory (RAM), an erasable programmable read-only memory (EPROM), an electrically erasable programmable read-only memory (EEPROM), a magnetic card or an optical card, a flash memory, or a tangible machine-readable memory used to transmit information through propagation signals (such as a carrier signal, an infrared signal and a digital signal) in electricity, light, sound or other forms through the Internet. Therefore, the machine-readable medium includes any type of machine-readable medium suitable for storing or transmitting an electronic instruction or information in a machine (such as the computer) readable form.

In the attached drawings, some structural or method features may be shown in a specific arrangement and/or order. However, it should be understood that such a

specific arrangement and/or ordering may not be required. Rather, in some embodiments, these features may be arranged in a manner and/or order different from that shown in the illustrative drawings. In addition, the inclusion of structural or method features in a particular figure does not imply that such features are required in all embodiments, and in some embodiments, these features may not be included or may be combined with other features.

It should be noted that the units/modules mentioned in the device embodiments of the present application are logical units/modules. Physically, a logical unit/module may be a physical unit/module or may be a part of a physical unit/module, or may be implemented as a combination of multiple physical units/modules. The physical implementation of these logical units/modules is not the most important. The combination of the functions implemented by these logical units/modules is the key to solving the technical problem presented in the present application. In addition, in order to highlight the innovation part of the present application, the above device embodiments of the present application do not introduce units/modules that are not closely related to solving the technical problems raised by the present application, and this does not mean that there are no other units/modules in the above device embodiments.

It should be noted that in the examples and the description of the present application, relational terms such as "first" and "second" are only used to distinguish one entity or operation from another entity or operation, and do not necessarily require or imply that there is any such actual relationship or order between these entities or operations. Moreover, the terms "comprise", "include" or any other variants thereof are intended to cover the non-exclusive inclusion, so that a process, method, system, product, or device that includes a list of elements not only includes those elements but also includes other elements that are not expressly listed, or further includes elements inherent to such a process, method, article, or device. An element defined by "includes a ..." does not, without more constraints, preclude the existence of additional identical elements in the process, method, product, or device that includes the element.

Although the present application has been illustrated and described by referring to certain preferred embodiments of the present application, those of ordinary skill in the art should understand that various changes can be made in form and details without departing from the spirit and scope of the present application.

Claims

1. An image super-resolution processing method applied to an electronic device, comprising:

obtaining a to-be-processed image; and

performing the following steps for the to-be-processed image in a n -level super-resolution processing network:

predicting a n^{th} -level blurred region in the to-be-processed image, and performing super-resolution processing on the n^{th} -level blurred region in the to-be-processed image to obtain a n^{th} -level super-resolution image;

predicting n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and $(n-m)^{\text{th}}$ -level blurred regions in a $(n-m)^{\text{th}}$ -level super-resolution image, and performing super-resolution processing on the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and $(n-m)^{\text{th}}$ -level blurred regions in the $(n-m)^{\text{th}}$ -level super-resolution image to obtain the $(n-m)^{\text{th}}$ -level super-resolution image, wherein a value of m is 1,..., $n-2$ in turn, to obtain a second-level super-resolution image; and

for the second-level super-resolution image, predicting the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred regions in the second-level super-resolution image, performing super-resolution processing on the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., first-level blurred regions in the first-level super-resolution image to obtain the first-level super-resolution image, and using the first-level super-resolution image as an output image;

wherein the degree of blur from the n^{th} -level blurred region to the first-level blurred region is reduced in turn, and both n and m are positive integers.

2. The method according to claim 1, wherein:

a difference between the n^{th} -level blurred region in the to-be-processed image and a n^{th} -level blurred region label image makes a value of a n^{th} region loss function in a corresponding predefined numerical range;

a difference between the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and $(n-m)^{\text{th}}$ -level blurred regions in the $(n-m)^{\text{th}}$ -level super-resolution image and a $(n-m)^{\text{th}}$ -level blurred region label image makes a value of a $(n-m)^{\text{th}}$ -level region loss function in a corresponding

predefined numerical range; and

a difference between the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred regions in the first-level super-resolution image and the first-level blurred region label image makes a value of a first region loss function in a corresponding pre-defined numerical range;

wherein the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred region label images are obtained by dividing the preset image according to n depth values ranging from large to small, and the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred region label images are all single-channel images having the same size as the to-be-processed image.

3. The method according to claim 2, wherein the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred region label images are obtained by using a DenseDepth network.

4. The method according to claim 1, wherein a difference between the n^{th} -level super-resolution image and a n^{th} -level blurred super-resolution label image makes a value of a n^{th} super-resolution loss function in a corresponding predefined numerical range;

a difference between the $(n-m)^{\text{th}}$ -level super-resolution image and a $(n-m)^{\text{th}}$ -level blurred super-resolution label image makes a value of a $(n-m)^{\text{th}}$ super-resolution loss function in a corresponding predefined numerical range; and

a difference between the first-level super-resolution image and a first-level blurred super-resolution label image makes a value of a first super-resolution loss function in a corresponding predefined numerical range;

wherein the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred super-resolution label images are obtained based on the fusion of samples in a preset data set, and each sample in the preset data set comprises a high-resolution image and n blurred images with different depth values, and the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., first-level blurred super-resolution label images have the same size as the to-be-processed image, and are all RGB multi-channel images.

5. The method according to claim 4, wherein before the obtaining a to-be-processed image, the method further comprises:

training the n-level super-resolution processing network by using the samples in the preset data set;

wherein in the training process, an input image of the n-level super-resolution processing network is obtained by processing the samples in the preset data set based on a target weight distribution relationship;

during the training process, the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred super-resolution label images are respectively obtained by processing the samples in the preset data set based on n^{th} , $(n-1)^{\text{th}}$,..., and first weight distribution relationships; and

each weight distribution relationship is used for one or more images to be fused in each sample, and a weight of each image to be fused; and the n^{th} , $(n-1)^{\text{th}}$,..., and first weight distribution relationships indicate that the smaller the depth value, the greater the weight of the blurred image.

6. The method according to claim 5, wherein:

when $n=3$, for a sample of a target high-resolution image, a first blurred image, a second blurred image, and a third blurred image comprised in the preset data set, the target weight relationship is used to indicate that the set output image comprises the following content: a first-level blurred region obtained by fusion of a weight α_1 *a target high-resolution image + a weight β_1 *the second blurred image, a second-level blurred region obtained by fusion of a weight α_2 *the first blurred image + a weight β_2 *the second blurred image; a third-level blurred region obtained by fusion of a weight α_3 *the second blurred image + a weight β_3 *the third blurred image;

a third weight relationship is used to indicate that a third-level blurred super-resolution label is set to comprise the following content: a first-level blurred region obtained by fusion of a weight α_1 *a target high-resolution image + a weight β_1 * the second blurred image, a second-level blurred region obtained by fusion of a weight α_2 *the first blurred image + a weight β_2 * the second blurred image; a third-level blurred region obtained by fusion of a weight α_4 *the second blurred image + a weight β_4 *the second blurred image;

a second weight relationship is used to indicate a second-level blurred

super-resolution label is set to comprise the following content: a first-level blurred region obtained by fusion of a weight α_1 *a target high-resolution image + a weight β_1 *the second blurred image, a second-level blurred region and a third-level blurred region obtained by fusion of α_5 *a target high-resolution image + a weight β_5 * the first-blurred image; and

a first weight relationship is used to indicate that a first-level blurred super-resolution label is set to comprise the following content: the whole image directly uses the target high-resolution image;

wherein the depth values of the first blurred image, the second blurred image, and the third blurred image sequentially increase.

7. An image super-resolution processing apparatus applied to an electronic device, comprising:

an obtaining module, configured to obtain a to-be-processed image; and

an execution module, configured to perform the following steps for the to-be-processed image obtained by the obtaining module in a n-level super-resolution processing network:

predict a n^{th} -level blurred region in the to-be-processed image, and perform super-resolution processing on the n^{th} -level blurred region in the to-be-processed image to obtain a n^{th} -level super-resolution image;

predict n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and $(n-m)^{\text{th}}$ -level blurred regions in a $(n-m)^{\text{th}}$ -level super-resolution image, and perform super-resolution processing on the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and $(n-m)^{\text{th}}$ -level blurred regions in the $(n-m)^{\text{th}}$ -level super-resolution image to obtain the $(n-m)^{\text{th}}$ -level super-resolution image, wherein a value of m is 1,..., n-2 in turn, to obtain a second-level super-resolution image; and

for the second-level super-resolution image, predict the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred regions in the second-level super-resolution image, perform super-resolution processing on the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., first-level blurred regions in the first-level super-resolution image to obtain the first-level super-resolution image, and use the first-level super-resolution image as an output image;

wherein the degree of blur from the n^{th} -level blurred region to the first-level

blurred region is reduced in turn, and both n and m are positive integers.

8. The apparatus according to claim 7, wherein

a difference between the n^{th} -level blurred region in the to-be-processed image and a n^{th} -level blurred region label image makes a value of a n^{th} region loss function in a corresponding predefined numerical range;

a difference between the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and $(n-m)^{\text{th}}$ -level blurred regions in the $(n-m)^{\text{th}}$ -level super-resolution image and a $(n-m)^{\text{th}}$ -level blurred region label image makes a value of a $(n-m)^{\text{th}}$ -level region loss function in a corresponding predefined numerical range; and

a difference between the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred regions in the first-level super-resolution image and the first-level blurred region label image makes a value of a first region loss function in a corresponding pre-defined numerical range;

wherein the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred region label images are obtained by dividing the preset image according to n depth values ranging from large to small, and the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred region label images are all single-channel images having the same size as the to-be-processed image.

9. The apparatus according to claim 8, wherein the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred region label images are obtained by using a DenseDepth network.

10. The apparatus according to claim 7, wherein a difference between the n^{th} -level super-resolution image and a n^{th} -level blurred super-resolution label image makes a value of a n^{th} super-resolution loss function in a corresponding predefined numerical range;

a difference between the $(n-m)^{\text{th}}$ -level super-resolution image and a $(n-m)^{\text{th}}$ -level blurred super-resolution label image makes a value of a $(n-m)^{\text{th}}$ super-resolution loss function in a corresponding predefined numerical range; and

a difference between the first-level super-resolution image and a first-level blurred super-resolution label image makes a value of a first super-resolution loss function in a corresponding predefined numerical range;

wherein the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred super-resolution label images are obtained based on the fusion of samples in a preset data set, and each sample in the preset data set comprises a high-resolution image and n blurred images with different depth values, and the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., first-level blurred super-resolution label images have the same size as the to-be-processed image, and are all RGB multi-channel images.

11. The apparatus according to claim 10, wherein the apparatus further comprises:

a training module, configured to: before the obtaining module obtains the to-be-processed image, train the n -level super-resolution processing network by using the samples in the preset data set;

wherein in the training process, an input image of the n -level super-resolution processing network is obtained by processing the samples in the preset data set based on a target weight distribution relationship;

during the training process, the n^{th} -level, $(n-1)^{\text{th}}$ -level,..., and first-level blurred super-resolution label images are respectively obtained by processing the samples in the preset data set based on n^{th} , $(n-1)^{\text{th}}$,..., and first weight distribution relationships; and

each weight distribution relationship is used for one or more images to be fused in each sample, and a weight of each image to be fused; and the n^{th} , $(n-1)^{\text{th}}$,..., and first weight distribution relationships indicate that the smaller the depth value, the greater the weight of the blurred image.

12. The apparatus according to claim 11, wherein

when $n=3$, for a sample of a target high-resolution image, a first blurred image, a second blurred image, and a third blurred image comprised in the preset data set, the target weight relationship is used to indicate that the set output image comprises the following content: a first-level blurred region obtained by fusion of a weight α_1 *a target high-resolution image + a weight β_1 *the second blurred image, a second-level blurred region obtained by fusion of a weight α_2 *the first blurred image + a weight β_2 *the second blurred image; a third-level blurred region obtained by fusion of a

weight α_3 *the second blurred image + a weight β_3 *the third blurred image;

a third weight relationship is used to indicate that a third-level blurred super-resolution label is set to comprise the following content: a first-level blurred region obtained by fusion of a weight α_1 *a target high-resolution image + a weight β_1 * the second blurred image, a second-level blurred region obtained by fusion of a weight α_2 *the first blurred image + a weight β_2 * the second blurred image; a third-level blurred region obtained by fusion of a weight α_4 *the second blurred image + a weight β_4 *the second blurred image;

a second weight relationship is used to indicate a second-level blurred super-resolution label is set to comprise the following content: a first-level blurred region obtained by fusion of a weight α_1 *a target high-resolution image + a weight β_1 *the second blurred image, a second-level blurred region and a third-level blurred region obtained by fusion of α_5 *a target high-resolution image + a weight β_5 * the first-blurred image; and

a first weight relationship is used to indicate that a first-level blurred super-resolution label is set to comprise the following content: the whole image directly uses the target high-resolution image;

wherein the depth values of the first blurred image, the second blurred image, and the third blurred image sequentially increase.

13. A readable medium, wherein an instruction is stored on the readable medium, and when the instruction is executed on an electronic device, the electronic device is caused to execute the image super-resolution processing method according to any one of claims 1 to 6.

14. An electronic device, comprising: a memory configured to store an instruction executed by one or more processors of the electronic device, and a processor, being one of the processors of the electronic device, and configured to perform the image super-resolution processing method according to any one of claims 1 to 6.