Photosounder

VERSION 1.8.2

User Guide

TABLE OF CONTENTS

I. Getting started

INSTALLATION

The demo version of Photosounder is simply installed by downloading it on the Photosounder website, and extrating the ZIP file to the desired location.

The full version is obtained by buying on the website. Once the payment has been received, you should receive by e-mail within 24 hours a link to download the program as well as the license file needed by the full version of the program to run. The full version is installed by extracting the ZIP file and placing the license file in the same directory as the program.

BASIC OPERATION

Photosounder can open many types of image and sound files. Whereas images are loaded largely untransformed, sounds first have to be processed and turned into a kind of image called spectrogram. That image represents features present in the sound loaded in a way that allows it to be turned back into sound in a relatively faithful way. A notable difference between images obtained in such a way and regular images in Photosounder is the former present a much higher precision in their intensity, and to represent this higher precision in a convenient fashion, their intensity is displayed in a very colourful way that helps you seize the slightest variations in the image in a glance.

Once either an image or sound is done loading (the latter which can take a few seconds depending on the length of the sound and the processing power of the machine it is running on), Photosounder starts synthesising the image into a new sound, which progress is depicted by the blue progress bar above the image. The light blue part of the bar is the part of the sound that remains to be synthesised, the medium blue area is the part of the sound currently being processed, and the dark blue part of the bar that displays a waveform is the sound you can already listen to by click the button Play. You can save the produced sound into a sound file by pressing the Save button. You can also save the image obtained by pressing the same Save button, and by choosing the image option in the drop-down menu. You can save and entire project (including every layer, setting and lossless data) by choosing to save to a Photosounder Archive file (.pha).

By default at the start, the Cursor tool is selected. It means that when you click somewhere on the image while playing the sound, the sound playback jumps where you clicked. You can also control the sound volume by turning the Volume knob in circles using the mouse. When the cursor is over the image, it will also indicate in the lower left corner of the window the time, frequency and pitch coordinates of where in the image your cursor is pointing. It can be particularly helpful if you're trying to determine the precise pitch of a note.

The window of Photosounder can be resized by clicking and holding the group of dots in the lower right corner of the window, and dragging the cursor to the position on the screen where the lower right corner of the window is desired to be.

CONTROLS OPERATION

In Photosounder, most settings are defined by knobs, which are controlled by the mouse. In order to change a value defined by a knob, simply click and hold anywhere on the circular area of the knob, and move the cursor in circles around the knob clockwise or counter-clockwise. For more precision in your movements, you can move the cursor further away from knob. For yet more precision, you can click and hold the right mouse button of your mouse, either simultaneously with the left button or independently, which allows you to spin the knob around 100 times slower.

PARAMETERS

A description of what each parameter means, from left to right and top to bottom :

Min. frequency	Lowest frequency in the image, at the bottom of it. By default 27.5 Hz, also know as A0, one of the lowest notes we can hear.
Max. frequency	Highest frequency in the image, at the top of it. By default 20 kHz, a frequency at the boundary of ultrasounds.
Frequency scale logarithmic base	Scaling of the frequency (vertical) axis going from linear (by convention here 1) to logarithmic (here defined as 2). Any setting in between 1 and 2 is an interpolation, a form of compromise, between linear and logarithmic scaling. The logarithmic scale is used by default, as it's closer to the way we perceive frequency, but it's also the same frequency scaling as used in musical notation, which explains why notes and octaves on the right side of the image are evenly spaced.
Volume	Volume of the playback sound. +0.0 dB means that the sound is normalised as to make the sound as loud as possible under any circumstance while avoiding clipping, which causes distortions. Therefore a volume above 0 dB should only be used to enhance the more quiet parts of a sound.
Time resolution	Playback speed of the image. By default set to 100 pixels per second, which means that a picture 800 pixels long will make a sound 8 seconds long. Also the rate of images obtained from sound, which means that if you open a sound and set the time resolution to 0.5 pixels/second, the sound will play 200 times slower than its normal rate, while if you set it to 2,000 pixels/second it will make the sound 20 times faster than usual, without changing the pitch of the sound.
Spray width	Width of the area affected by the spray tools in pixels, in the scale of the screen.
Tool intensity	Intensity of the spray tools or rectangle tool in percents.
Rotation	Rotates the image by the indicated angles, always with respect to the original orientation.
Horizontal flip (FLIP)	Reverses the image in time (horizontally). Combined with a rotation by 90° or 270° it can also be used to reverse an image in frequency (vertically).
Invert (INV.)	Inverts the colours/intensity of an image in the same way as a photographic negative.
Mask Invert (M.I.)	Inverts the mask. Produces the difference between the original image and the current one.
Vision Mode (V.M.)	Switches the Vision Mode.
Gamma	Gamma function, as it can be found in most image editing programs. Can be thought of as a sort of brightness function. The default (and neutral) value is defined here to be 1:1.
Layer intensity	Changes the intensity of a layer with respect to other layers.

SPRAY TOOLS

In addition to the various effects you can achieve by adjusting the aforementioned parameters, you can also finely edit the features of an image-sound by using the Spray tools. The Spray tools, activated by clicking the second, third or fourth icons to the left of the image, can be used to highlight, tone down or erase features from image-sounds, or painting up new ones. It can be used to achieve many things, such as singling out a sound, suppressing a sound, reproducing traditional sound effects by imitating their effects in the time-frequency domain (this can be learnt by loading a sound and a copy of it with the desired effect applied, and looking at the visual differences between the two), or even creating sounds from scratch.

Using these tools, a left click on the image will darken, brighten, paint up or smart erase (depending whether you've selected the black, white, flashlight or road roller spray tool) the area in the vicinity of the mouse cursor, area which will be defined by the Spray width parameter, with an intensity controlled by the Tool intensity setting and by how long the click is held. Inversely, if the right click is held, the same area will be progressively reverted to its original state. Therefore anything done with the Spray tools can be easily reverted. The reason for this reversibility is that a copy of the original image is kept in memory, and is used for reverting changes.

RECTANGLE TOOL

The rectangle tool works similarly as spray tools. Holding the left mouse button darkens the chosen area by a ratio defined by the Tool intensity knob, holding the right mouse button inversely brightens the area.

HARMONICS MODIFIER

If toggled, the harmonics modifier, represented by four vertically stacked dots, modifies the behavior of all spray tools by reproducing their action on the harmonics of the frequency that the mouse cursor points to. This allows for the simultaneous modification, suppression or creation of harmonic content.

MAGNET MODIFIER

If toggled, the magnet modifier snaps the mouse cursor to the nearest feature, determined by proximity and intensity. If the harmonics modifier is also toggled, the magnet modifier also takes into account harmonic content to determine which point to snap to. The point on the image to which the magnet modifier snaps is represented by a diagonal cross on the screen.

ZOOM

Photosounder allows you to zoom in on the image you're editing using the scroll wheel of your mouse. Scrolling up will zoom in in time, scrolling down will zoom out. You can zoom in frequency instead by holding the Shift key while scrolling, or zoom in both horizontally and vertically by holding the Control key. Once zoomed in, you can drag scroll the image by holding the middle mouse button or the End key and mouse the cursor around. Note that if you do not have a scroll wheel on your mouse you can use the Page Up and Page Down keys instead.

II. Lossless mode

DESCRIPTION

By default, Photosounder synthesises sounds purely from image data, discarding all original input sounds. While powerful and offering much flexibility, this approach also introduces a loss in sound quality. The lossless mode remedies to that by using the original sound as a basis for resynthesis, in the place of noise. It works by keeping an original sound and its untouched image, then comparing the differences between the image as modified by the user and that original image, and applying the differences to the original sound. Therefore, as long as a sound's image is untouched, the resulting sound will be strictly similar to the original sound, hence the term lossless.

USAGE

In lossless mode, the last sound loaded and its associated original image are always kept in memory, and used in combination with the current image to produce the synthesised sound. This means that you can open a sound, export its image, edit this image in an image editor, then open this modified image with Photosounder in lossless mode, the resulting sound being the differences between the new image and the original image applied to the original sound.

The lossless mode is recommended for tasks such as erasing, filtering, isolating or enhancing sound features, noise reduction, subtraction and other operations between two different sounds and various time-frequency mastering operations.

LIMITATIONS

However, there are limitations inherent to the lossless mode which are important to keep in mind. The lossless mode works by comparing the matching pixels of the original sound's image and the current image and filtering the original sound in time-frequency according to those differences. This means that it is important not to displace anything in either time of frequency in order obtained the desired quality of sound. This is why several controls are locked in lossless mode, as using them would in the best cases produce a result of quality possibly worse than in lossy mode. It is also recommended to keep images to their original length.

III. Layers

USAGE

Layers constitute a stack of individual images that combined together make up the final image that you see and hear. Using multiple layers allows you to do things you wouldn't be able to do otherwise.

By default when you open a sound, image or create a new image only one layer is present. You can add a new blank layer on top of that layer by selecting New Layer in the Layers menu, or duplicate that layer by selecting Duplicate Layer. You can see the list of layers at the bottom of the Layers menu. The currently selected layer is indicated by a check mark next to its name. You can delete the currently selected layer by clicking Delete Layer in the Layers menu. You can also move the currently selected layer up and down in the stack of layers by selecting Move Layer Up or Move Layer Down. You can switch between a view of all the layers combined or a view of only the currently selected layer by clicking View Current Layer Only.

Each layer has a blending mode associated with itself. The default mode is Addition, you can see in the Layers menu that it is already checked. The blending mode defines how each layer is affects the combination of the layers below. Separators can be added between layers to create isolated groups of layers. Layers between separators are combined separately from other groups of layers. The combinations of each group of layers are then combined together according to the blending mode set for each separator. This can for example be used so that a layer with a blending mode set to Multiplication would only affect a precise group of layers and not every single layer below it.

Layers can also be merged into one using the Merge Down function in the Layers menu. The Merge Down function merges the currently selected layer with the layer right below it. For this reason, the proper way to merge a group of several layers is to start by selecting the second layer from the bottom and to use Merge Down until all are merged together. Merging layers together in a different order could yield a result inconsistent with what was seen prior to the merging depending on the blending modes used. The combination of all layers can also be turned into a whole new layer at the top of the stack of

layers using the Merge All to New Layer function.

IV. Keyboard shortcuts

For the sake of convenience for the experienced user, or for the sake of keeping the visual interface simple and practical, some functions of Photosounder are controlled by keyboard shortcuts. Some of these shortcuts make it faster to access functions already represented visually, while some give access to practical functions which are invisible on the interface.

VISIBLE FUNCTIONS

O	Equivalent of the **Open** button
S	Equivalent of the **Save** button
Esc	Equivalent of the **Quit** button
Space	Equivalent of the **Play/Pause** button

INVISIBLE FUNCTIONS

C	**Copy** function. Copies the current image to the clipboard, allowing it to be pasted in other image editing programs.
R	**Reload** function. Reloads the sound or image from the file. Most practical when modifying in parallel the input file with another image or sound editing program to quickly reload a file freshly updated from another program.
U	**Undo** function. Cancels the last action on the image. Redoes it if pressed again.
E	**Erase changes** function. Cancels everything that has been done with the Spray tools or Rectangle tool from the beginning.
D	**Default values** function. Resets the default values for the knobs Min./Max. Frequency, Time resolution, Gamma, Logarithmic base, Spray width and intensity.
N	**New layer**. Adds a new layer on top of the selected one.
J	**Joins** the selected layer with the layer right under it.
Del	**Deletes** the selected layer.
V	**View** current layer only switch.
L	**Loop** function. Loops back to the beginning of the sound if the key is held when the end of the sound is reached.
B	**Backwards** playback function. Plays the sound backwards when held.
L + B	**Ping-pong** playback function. Plays the sound back and forth when the two are simultaneously held. If B is pressed before L then it will start by playing the sound backwards, until it reaches the beginning of the sound and starts playing forwards again.
H	**High precision cursor** function. When held the mouse cursor goes at a

	fraction of its normal speed as to allow high precision editing.
F	**Force** function. When held, this button allows you to force the modification of locked (grayed out) controls when in lossless mode.
Down Arrow	**Full rewind** function. Jumps back to the very beginning of the sound.
Up Arrow	**Stop** function. Jumps to the very end of the sound.
Left Arrow	**Jump back** function. Rewinds the sound by 1 second.
Right Arrow	**Jump forward** function. Fast forwards the sound by 1 second.
Scroll Up / Page Up	**Zoom in** horizontally.
Scroll Dn / Page Dn	**Zoom out** horizontally.
Shift + Scroll	**Zoom** in/out **vertically**.
Ctrl + Scroll	**Zoom** in/out both ways.
Middle Click / End	**Drag scroll** the image.

V. Advanced options

STEREO OPERATIONS

Photosounder operates entirely in mono. However, there are four options in the Options menu that can be used to treat each channel of a stereo sound separately and join the output into a stereo sound file.

The **Load Left Channel Only** and **Load Right Channel Only** options can be used so that when a stereo sound is loaded one of the two channels is loaded instead of an even mono mix of the two. You can therefore treat each channel separately and save the results for each into a separate sound file. However, before you save, you should make sure to toggle the **Unnormalised Output** option. Usually sounds are normalised so that their highest peak is set to 0 dB. This means that in the process of normalisation, one of the two channels might be amplify more than the other, therefore offsetting the pan in either direction. This does not occur when Unnormalised Output is enabled. However note that if one channel appears to be 100 times (+40 dB) louder than the other, just amplify the other channel accordingly by +40 dB. Unnormalised Output amplifies the output by multiples of 100 (40 dB) so that the peak amplitude of the output is always between -40 dB and 0 dB.

Once you've saved both the left and right channel to individual files, you can join them together into one stereo file by using **Mix Channel Files to Stereo...** in the Options menu. A first open dialog will appear with which you should select the left channel file. A second open dialog will appear so that you can select the right channel file. Then a save dialog will appear so that you can save the mix of the two files to a new sound file.

MISCELLANEOUS OPTIONS

In the Options menu, **Loop Playback** can be toggled so that the sound playback always loops.
Force Fast Image Display can be enabled so that the fastest method for displaying the image on the screen is used. Likewise, **Force Smooth Image Display** can be used for the slower but smoother method for displaying the image to be used at all times. When neither of the two are enabled Photosounder uses either of the two methods depending on the context. **Long Block Duration** can be enabled so that synthesis uses a noise that loops every 60 seconds instead of one that loops every 2 seconds. This is recommended in cases the 2 second repetition is undesirable.

CONFIGURATION

Some parameters used by Photosounder can be permanently altered by modified the file 'config.txt' created by Photosounder in its folder. The numerical values can be changed by editing this file in a text editor, the changes taking effect the next time Photosounder is started. Note that the configuration file is regenerated every time Photosounder starts to reflect how the contents of the file were interpreted.

audio_buffer_size	Size of the audio buffer used for audio playback in samples. If you experience stutter you should try to increase it to such values as 4096, 8192 or even 16384. Otherwise you can try lowering it to other powers of two such as 512 or 1024 in order to obtain less delay in the sound playback.
audio_sample_rate	Internal sampling rate used playback, internal processing and audio output.
waveform_display_height	Height of the waveform display in pixels. An odd height is preferred.
cursor_slow_rate	Fraction of the normal mouse cursor speed when the H key is held down.
spray_harmonics	Number of harmonics (base frequency included) the harmonic modifier creates.
min_frequency	Minimum frequency used for analysis and by default.
max_frequency	Maximum frequency used for analysis and by default.
pixels_per_second	Time resolution in pixels/second used for analysis and by default.
bands_per_octave	Vertical resolution of the image after analysis defined by pixels/octave.
min_bpo	Minimum effective pitch resolution in pixels/octave used in the analysis.
max_time_res_for_min_bpo	Time resolution in pixels/second at which the time resolution will stop increasing while the pitch resolution increases from the minimum pitch resolution defined by **min_bpo**.

The last four parameters define the varying effective time and pitch resolutions in the following manner. The pitch resolution in the lowest frequencies is defined by **min_bpo**. As we go up in pitch, the time resolution increases, while the pitch resolution remains constant, until we reach the time resolution defined by **max_time_res_for_min_bpo**. Then the time resolution remains at that value while the pitch resolution increases. When the pitch resolution reaches **bands_per_octave**, the time resolution starts increasing again until it reaches **pixels_per_second**. The first two parameters are independent from the two others, which means that you can safely increase **pixels_per_second** and **bands_per_octave** without the appearance of the result being affected for the lower frequencies. Likewise you can change the values of **min_bpo** and **max_time_res_for_min_bpo** to change the effective resolutions in the lower frequencies without having any impact on the results in the higher frequencies.

VI. Miscellaneous

MP3 SUPPORT

On Windows the loading of MP3 files cannot be supported by default for legal reasons. However you can very simply add the possibility for Photosounder to load MP3 files by downloading the file libmpg123-0.dll and copying it to the Photosounder folder, or optionally, your Windows installation's system32 folder.

This file can be downloaded at the website of [library mpg123](), the latest version as of the time of this writing being available [here]().

On Mac OS X, MP3 decoding is performed using the system's built-in frameworks, and therefore doesn't necessitate any third party addition.