Stereoscopic image quality control is more complex than mere 2D monitoring. You want to look at the stereoscopic disparities, more than the image in volume. You want to make sure that all the bad disparities, the vertical ones, are either nonexistent or within the required range. Regarding the good disparities, the horizontal ones, you want to make sure they remain, at the very least, in the comfort envelope. Further inspection will ensure that they match the artistic intent, or depth script. The ultimate judgment should be based on the depth feeling your imagery is gathering. And that’s the only one that requires actual 3D vision.

How does the progression of quality assurance in depth acquisition occur? Its science and procedures are presented here in three sections. We’ll distinguish

- Image monitoring, a visual evaluation using basic signal formatting.
- Image analysis, an automatic extraction of the stereoscopic descriptors of a picture.
- Image correction, the resultant grooming of a stereoscopic picture.

We built this distinction on our observation of the wide range of implementation complexity and cost of such systems.

*Image processing* involves only basic 2D image manipulation, like 50% mixes or frame-compatible formatting. Its typical implementation is in stand-alone conversion boxes called *stereoscopic multiplexers* or inclusion in a 3D display. Its application domain is the image monitoring.

*Image analysis* refers to complex processes involving software, computers, multi-image processing, and, in the most advanced cases, full 3D reconstruction. Some image analyzers provide a feedback loop to motorized rigs.

*Image correction* can be via a feedback loop on the rig’s motion control, or in the image domain, in real time or in batch processing. The latter is a software solution that would run overnight on your laptop, and almost instantly on a render farm. The former may be the very same program, but implemented on
dedicated processing hardware, with the ability to correct your live 3D stream as you shoot it, even before it reaches the OB van grid.

**MONITORING 3D**

The main rule is that you want to see the 3D. Clearly, you need to monitor your work in 3D. You wouldn’t monitor a color shoot on a black and white monitor, nor would you record sound without a headset, just by watching the levels. For the same reasons you wouldn’t do these things, you shouldn’t shoot 3D without being able to see it.

That being said, this does not mean it has to be in stereoscopic 3D. To start with, not everybody needs to monitor the 3D on a 3D display, and there’s even a debate on the importance of running dailies in theater to check 3D in its final screen size. As we’ll see, the best monitoring for stereoscopic images is sometimes on a 2D display. This requires the images to be adequately formatted into a structure that will reveal their potential defaults.

**Stereoscopic Image Processing for 3D Monitoring**

Once again, let’s refer back to our color analogy. There are three levels of color monitoring: basic control is watching lowlights and highlights, making sure your images are not burned out; intermediate color control just requires an all-purpose monitor; and advanced color control requires a color-calibrated monitor, light-controlled environment, or the trained eyes of an expert colorist or DP. Otherwise, reference patterns and a vectorscope would do the job, even with a colorblind technician. Beyond this, onset LUT management will allow the crew to make sure the tweaked image will match the desired artistic effect.

**BASIC 3D MONITORING**

Basic stereoscopic control can be done with a 50% mix on a regular 2D monitor. Vertical disparities should be unnoticeable on a small field monitor; otherwise they’ll be uncomfortable on full-size screens.

Side-by-side and over/under modes can also be used, but they are less efficient in detecting horizontal or vertical disparities and rotations. Anaglyphic 3D is often preferred for previews and is best enjoyed without glasses. The red and blue fringes act as disparity enhancers on the display.

**INTERMEDIATE 3D MONITORING**

Intermediate level quality control is done using a difference view that clearly shows the disparities and makes them much more noticeable, even measurable.

Basic and average controls can be supplemented with markers showing reference parallaxes on the monitor. One can use simple Post-its or tape applied on the screen. Some professional displays can electronically insert visual helpers,
like gridlines, bars, or checkerboards. If you are shooting with no target screen size, 1% and 2% of the screen width are the regular landmarks. If you are shooting for a defined screen size, the size of the reference grid unit can be the equivalent of one or two human inter-optical distances. On a feature, the
reference maximum positive parallax will be chosen during the preproduction. Obviously, specific parallax references for FX shots will be enforced with even more caution.

**ADVANCED 3D MONITORING**

Advanced controls for specific image default detection—like rotations, keystones, and various geometry mismatches—can be performed with the naked eye and years of experience, or with specialized tools. Alignment patterns are one option. Image analysis is another. Combining the two is a great approach.

More complex image processing can be performed on a stereoscopic feed, without going all the way to the complexity of a full 3D analysis. Digital disk recorders built around a computer offer some advanced monitoring tools, such as

- Edge detection, for focus matching
- Interactive digital zoom, for fine alignment and convergence
- Alignment grid overlays, for parallax control
- False color zebras
- Spot meters
- Dual histograms, for color balance
- Parallax shifts, to adequately monitor parallel footage
- Anaglyph mixing, to be used with and without glasses
- Wiggle displays, to quickly switch between views

**Identifying 3D Image Defaults**

Most 3D-related headaches come from bad 3D images, a generic term under which stereographers include all pictures that suffered, at any stage of their life, from lack of caution, knowledge, experience, or luck. Mistakes occur very easily in 3D production and should be noticed and corrected at once, for some of them may be extremely costly to fix later on in the production process. This is one of the reasons why you should always check your work in 3D, and on the biggest screen available.

You may be feeling that this “watch your 3D, every time you can, and on a big screen” advice is printed on every other page of this book. That may be the case, and it would be a good thing. There are probably fewer than five directors or DPs on earth who are able to put together a 3D shot without needing to check it in 3D—but even these geniuses would never do such a thing; they will always visually check the 3D, one way or another.

In the next sections, fixes and tricks are sometimes suggested for low-end configurations using prosumer cameras. Obviously, the readers using high-end rigs and digital cinema cameras are invited to refer to their manuals to find the proper course of action to apply the proper fix to the image default.
FIGURE 3.3
The Technicolor Stereoscopic Certifi3D Chart. Another visual representation of the most common stereoscopic defaults. Courtesy of Technicolor.
INTER-OPTICAL DISTANCE

Cause: The distance between the cameras is not adequate.

Consequence: In cases of excessive IO, the 3D effect is too strong: the close-up elements are too close AND the far away elements are too far away. In cases of insufficient IO, the 3D effect is too weak: the difference between close-up and far away elements is too small, and the overall scene looks shallow. If 3D is too strong, it will prevent the audience from enjoying it, generating excessive positive and negative parallaxes, painful divergence, and toe-in.

Catch it: This is the reason why on-set 3D monitoring is a must. To make it safer and easier, compute your absolute maximum parallax in screen percentage, based on your final screen size. Then compute your monitoring screen relative maximum parallax, in inches or centimeters, based on its resolution and size. Eventually, draw a rule on a strip of adhesive tape and overlay it on the picture. Check it, or have it checked as often as needed. If you are using a computer-based recording or monitoring system, you can get the same results using digital overlays.

Correct it on set: On set, bring the camera to the right interaxial. If you have reached the limits of your rig, change the camera placement, focal length, or consider changing the rig. Get a beam-splitter rig for smaller inter-optical distances, or get a wider base on a parallel rig.

Cure it in post: Inter-optical errors are the most expensive to correct, for one of the eyes has to be regenerated in CG using 2D/3D conversions, or view-synthesis techniques.

CONVERGENCE

Cause: Cameras were converged on the wrong point, or one eye was excessively horizontally shifted.

FIGURE 3.4
Simulated IOD Default.
Image courtesy of Steve Schklair, 3ality Digital.
Consequence: The whole scene is either too far away or too close to the viewer.

Catch it: As usual, you can catch this by visually checking your 3D, and, if needed, relying on the visual helpers you draw on the monitoring screen.

Correct it on set: Turn the convergence knob on the rig. If you have reached its limits, you are shooting under a very strange configuration. Please send me a picture of the set. If you are sure you need more convergence, shoot with a huge over-scan, and you’ll reconverge in post.

Cure it in post: This is the easiest, cheapest, and most common 3D correction. It’s actually a mundane task in 3D to reconverge footage, from on-set to final grading and packaging. A 3D TV remote should even include a reconvergence button to let viewers fine-tune the image to their preferences, which is done by shifting one or two eyes sideways. However, this will most likely bring the image edge inside the visible frame, creating an undesired floating window that virtually moves the perceived screen in or out the theater space. In order to avoid this, a slight zoom in and crop should be applied to both eyes.

VERTEICAL ALIGNMENT

Cause: On set, one camera is most likely pointing up or down, or a zoom is off-axis. In post, one eye has been vertically shifted.

Consequence: One of the views is uniformly higher or lower than the other. This affects the readability of the 3D, not its strength or depth placement.
Catch it: During monitoring, apply horizontal stripes of tape or show guides in your digital recorder or compositing GUI. Check that left and right images of objects, shown in anaglyph, 50% mix or difference, are horizontally aligned.

Correct it on set: If the vertical parallax is more than half a percent of the picture on a high-end rig, the camera’s geometry should be fixed. On low-end and makeshift rigs, it may be useless to spend too much time fixing it. This glitch is easy to fix in post, at the cost of a slight zoom and crop of both eyes.

Cure it in post: Another easy fix: shift the images, just as in a convergence correction, but this time along the vertical axis. Zoom and crop are to be expected.

**FIGURE 3.6**
Simulated Vertical Alignment Default.
*Image courtesy of Steve Schklair, 3ality Digital.*

**KEYSTONE**

**Cause:** Keystone appears when the optical axes are not parallel, due to convergence or, less often, strong vertical misalignment.

**Consequence:** As seen in Figure 3.7, which presents the effects of toe-in convergence, keystone generates asymmetric zooming in the sides of the pictures, generating vertical misalignment in the four corners. According to stereographic director Brian Gardner, minor keystone is acceptable and may even help in perceiving 3D, to the extent that our visual system, being itself a toed-in converged system, experiences and uses keystone as a natural binocular vision byproduct. This unconventional take on 3D obviously applies to infinitesimal deformations, not to the huge and painful image distortions that have to be fixed.
Catch it: Using the 50% mix and the horizontal lines, you’ll see a picture that is okay in the center, but is symmetrically skewed in the corners. The top and bottom of the center third of the picture are okay.

Correct it on set: Maybe you should not converge on that shot. Shoot parallel, get a huge over-scan, and set the convergence in post.

Cure it in post: Fixing a keystone in postproduction is not an easy task, but a small correction is easy to set up and will be efficient. Apply a perspective distortion in the direction of the offending keystone, and then zoom and crop as usual. A significant amount of keystone requires complex depth warping and falls into the expensive fixes one should avoid. For shots that will be mixed with CG images, stereoscopic supervisors tend to prefer to work with footage shot parallel and then reconverge in postproduction.

FIGURE 3.7
Simulated Keystone Default.
Image by Bernard Mendiburu, based on Steve Schklair, 3ality Digital.

ROTATION

Cause: Image rotation appears when the camera’s optical axis is rotated along the Z axis.

Consequence: Horizontal alignment is progressively lost from the center to the four corners of the image. It generates painful vertical parallaxes on the sides and depth artifact in the center top and bottom areas.

Catch it: On the monitoring screen, the image is okay in the center and gets worse as you approach the edges. All four edges of the frame are off, in asymmetrical directions.
Correct it on set: There's something wrong with the rig, most likely a rigidity issue on a makeshift rig. If it's significant, it may be the sign of a mechanical default in the rig mechanical integrity, and therefore should be addressed as a safety warning. Sometimes, tension on a cable or momentum on an attachment changes the balance of one camera.

Cure it in post: Image rotation can efficiently be fixed in post, at the cost of a crop and resizing.

FIGURE 3.8
Simulated Rotation Default.
Image courtesy of Steve Schklair, 3ality Digital.

FOCAL LENGTH

Cause: Prime lenses are not always matched and zoom lenses are almost never matched, and therefore it is difficult to get two cameras zooming symmetrically. If you zoomed with a low-end rig using prosumer cameras, you will likely have to fix it in post.

Consequence: One image is bigger than another, the center is okay, and the 3D gets worse at the edges, with asymmetrical depth artifacts in the mid-high left and right areas.

Catch it: The 50% mix shows a picture matching in the center but not on the edges. The misalignments are radially symmetrical, from the image center into the four directions.
Correct it on set: Correct the focal length on one camera. If you are using a low-end camera with motor-only control of the zoom, you’d better save the trouble and plan to fix it in post.

Cure it in post: This is another glitch that’s easy to cure in post. Because most other corrections include a resize of the images, focal length adjustment usually comes for free in the touch-up package.

FOCUS MISMATCH

Cause: Low-end cameras have poor manual control on focus distance, and autofocus modes may disagree on the subject distance. Despite the efficiency of our visual system at pasting sharpness from one eye to the other, we eventually notice it when focus asymmetry gets too big. In CG imagery, it sometimes happens that one eye has no motion blur pass. This causes a strange feeling and adversely affects the fusion of the images. Focus matching is a visual quality element to be taken care of in 3D. The relation between the focus distance and focal length with zooms is another point of concern.

Consequence: One image is sharp, the other is blurry. Only certain depth ranges of the pictures are affected.

Catch it: Focus asymmetry is really hard to catch on a 50% mix and is better seen on a side-by-side view. On a 3D display, you will have to alternately close each eye and compare the images.
Correct it on set: High-end rigs with professional lenses and remote focus should never give you this trouble. On low-end rigs, this is much more common, and there’s no real cure if your cameras have no focus ring. In this case, you have one more reason to get more light and an infinite focus rather than a shallow focus.

Cure it in post: Obviously, this has to be taken care of on set. In postproduction it is close to impossible to regenerate sharpness; the sharper eye would have to be blurred to match the other one. A 2D/3D conversion may be the only solution.

**FIGURE 3.10**
Simulated Focus Mismatch Default. Looking at known sharp textures like human faces helps detect focus mismatch.

*Image courtesy of Steve Schklair, 3ality Digital.*

**DESYNCHRONIZED STEREO**

Cause: The two cameras are not running in perfect synchronism. If both left and right images are shot without paying great attention to time synchronization, issues will definitely occur. The faster the camera or the action moves, the higher the timing requirements. If your image is motion blurred, you need an extremely accurate genlock, down to the line or pixel clock, or the motion blurs will have vertical parallaxes.

Consequence: The slightest time delay will transpose camera and actor movements into unwanted parallaxes. If you took care of every mismatch listed earlier in this section, but missed the time synchronization, you will get shots with all the defects you tried so hard to control.
Catch it: Objects that have a cyclic movement, like a waving hand, are the best cues for incorrect timing. They show inconsistent parallax; the action in one picture is constantly catching up with the other.

Correct it on set: Check this with the director of photography, the camera technician, or the first, second, or third assistant—whoever is in charge of the genlock. If you are the person in charge, check the clock distribution and check that everything is perfectly symmetrical, down to the quality and the length of the cables.

There are some cameras that can be genlocked at video frame rates, like 25 or 30fps, but not at the digital cinema 24fps. The Sony PMW-EX3 has this reputation.

Cure it in post: Try retiming software, or some time warp filters from your favorite FX suite. While you are computing optical flows, you may give a chance to 2D/3D conversion too.

ROLLING SHUTTER ORIENTATION ERROR

Cause: You shoot some fast action with a mirror rig and the rolling shutter camera in the reflective position is not upside down.

Consequence: Both cameras are scanning the vertical axis in opposite directions, and the frame delay is up to one full frame from the top to bottom of the picture.

Catch it: This one is quite hard to catch visually. Pan the camera laterally and see the horizontal disparities build up asymmetrically between the top and the bottom of the picture. Spin a scale model windmill in front of the camera and see the wings bend toward or backward upon the rotation direction. You don’t
have your scale model windmill with you? The director is likely searching for you with a baseball bat. Try to secure it and have someone swing in front of the camera. That should do the trick.

Correct it on set: Turn one of the cameras upside down.

Cure it in post: Same as synchronization mismatch, besides the fact that you need to apply a vertical gradient to the retiming. However, we are not aware of such a feature in correction tools.

Beware: Active display of passive stereo and vice versa can interfere with this issue.

FIGURES 3.12A AND 3.12B
Images by Bernard Mendiburu, based on Steve Schklair, 3ality Digital.

ERRORS INDUCED BY THE HALF-MIRROR, AND OTHER ERRORS
The half-mirror in the rig may not be perfectly 50/50. In this case, some image asymmetry can be the direct or indirect effect of the asymmetry of the light beams.

Depth of field mismatch: Asymmetrical depth of field is generated by unequal apertures. First, check the camera settings: the gain, the shutter speed, the ND filters. On a low-budget production, the second possibility is that the cameras are on “auto” and the computers figured out two different sets of numbers. In any case, depth-of-field asymmetry will then come with lightness, contrast, and motion blur asymmetries.

Colorimetric mismatch: Colorimetric asymmetry is a staple default of the half-mirror. Even on a clean side-by-side rig, you’ll have a hard time matching the white balances. Our brain will do a great job at hiding such color mismatch. When you’re setting on a white reference, control that the camera agrees on the color temperature. Advanced cameras have complex presets, sometimes on memory cards. Make sure they match too.

Lightness mismatch: Uneven exposure will generate asymmetries and give or take texture in the highlights or shadows, creating an unpleasant binocular High Dynamic Range (HDR) experience. The interesting thing with lightness asymmetry
and 3D, is that it actually generates a time delay in the obscured eye. This delay will generate unwanted parallaxes in every direction the camera or subjects would be moving. On a dolly, or horizontal pan, it’ll tweak the depth. On action or vertical camera movement, it will affect the 3D perception.

3D Monitoring Tools

We are turning a corner, moving from makeshift 3D monitoring into professional gear. Not long ago, monitoring 3D required processing left and right images with a dedicated computer and plugging it in on one of the very few 3D displays available. Rear projection DLP TV came first, and passive polarized monitors followed more recently. The situation is now totally different with the following new products:

- Consumer 3D displays with no dual input and image processing.
- Professional 3D displays with integrated dual input and some 3D processing.
- Simple 3D multiplexers that get two images into a 3D format.
- Complex 3D processors with some image-manipulation capabilities.

CONSUMER 3D DISPLAYS

Consumer 3D displays have the ability to show 3D, but they were not designed to be used on set or on location. They cannot run on batteries, and most likely require the 3D images to be processed for display. They all cost from $1000 to $5000, with the plasmas, full HD projectors, and passive TVs in the upper margin.

If you are using these displays, remember to thoroughly investigate the setup menus to disable any image enhancement features like smooth motion and other cinema colors. The frame interpolation, which adapts the 24 fps input into the display’s 120 Hz refresh rate, is a known culprit. It is usually built around a modernized 3:2 pull-down to generate a pair of 60p video streams.

<table>
<thead>
<tr>
<th>Display</th>
<th>Input Format</th>
<th>Glasses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rear projection DLP TVs</td>
<td>Checkerboard FC</td>
<td>Actives with DLP link or IR with transmitter</td>
</tr>
<tr>
<td>3D-compatible projectors</td>
<td>120Hz active stereo</td>
<td>Actives with DLP link or IR with transmitter</td>
</tr>
<tr>
<td>Active 3D TV like plasma</td>
<td>Most FC formats</td>
<td>Actives, brand-compatible</td>
</tr>
<tr>
<td>Passive 3D TV</td>
<td>Most FC formats</td>
<td>Passives, most likely compatible with RealD cinema glasses</td>
</tr>
</tbody>
</table>
**PROFESSIONALS 3D DISPLAYS**

Professional 3D displays were specially designed for use on set and on location. Most of them run on batteries and have left and right input and SDI and HDMI connectors. Some have image-processing features for monitoring 3D, such as 50% mix, left/right difference, and reconvergence. Some monitors offer flip and flop image inversion to monitor 3D coming live from a mirror rig. Another useful function is an overlaid grid with variable pitch that allows for a quick check of excessive disparities.

Transvideo was the first company to offer a field monitor. The CineMonitorHD12 3DView is a classic 12-inch with left and right SDI inputs, and anaglyphic and active stereo modes.

The second generation of professional 3D field monitors use passive 3D displays, with polarizing filters that rotate each line in opposite directions. One can look at the 3D picture with low-end disposable 3D cinema glasses, or with high-end designer glasses that offer a much better contrast.

Panasonic has a 25.5-inch 3D monitor, the BT-3DL2550, with dual SDI and DVI-D inputs and a 1920x1200 resolution. It displays 720p, 1080i, and 1080p video in 10 bits color depth.

Panasonic plasmas are to be considered too. They offer full resolution, thanks to active 3D, high brightness, saturation, and contrast. The stereo separation is great, with no ghosting or sweet spot effect. Professional models for field use are not yet in catalogs.

Sony offers the LMD2451TD, a 24-inch monitor with the same 1920 × 1200 resolution. The LMD4251TD is a 42-inch version.

JVC has a 46-inch monitor, the GD-463D10U, and at NAB 2009 presented a 4 K 3D monitor prototype.

Assembling two monitors in a beam-splitter mount will give you a full-resolution passive display. Cinetronic, among many others, offers such a system. It was used in the production of *Pirates of the Caribbean 4*.

**3D MULTIPLEXERS AND IMAGE PROCESSORS**

All 3D multiplexers are signal converters used to convert the image stream initiated at the camera into a format that a 3D display or a 3D recorder will understand and process adequately. Their main uses include:

- Getting the signal from a pair of (DS) 2D cameras into a (FC) 3D monitoring or recording device.
- Converting bit streams between consumer and professional connectivity standards.
Retrofitting a legacy 2008 3D-ready display into a modern 2010 3D-capable device.

Bit stream conversions include:

- Converting dual-stream (DS) 3D into frame-compatible (FC) 3D, with two inputs and one output.
- Converting FC 3D into DS 3D, with one input and two outputs.
- Digital format conversion, to and from DVI, HDMI, and HD-SDI.
- Active/passive stereo conversion to and from 120 and 144 fps.
- Sometimes digital/analog conversion, DVI/VGA.

Image manipulations include:

- Image axis permutation, called flip or flop, for mirror rig live monitoring.
- Horizontal image translation (HIT), used when shooting parallel material.
- 50% mix, left/right difference, and anaglyph image encoding for basic 3D monitoring.

There are several types of models:

- Miranda DVI-Ramp 2 (a dual-channel genlockable DVI/SDI converter)
- Cine-tal Davio
- Inition StereoBrain
- Doremi Dimension3D (designed to connect PCs to DCI projectors)
- AJA Hi5-3D, aka the Keith Box (a pocket-sized camera companion device)
- Matrox (developing a 3D multiplexer with Tim Dashwood for NAB 2011)
- Blackmagic Design (introduced its conversion box, the HDLink Pro 3D at IBC 2010)
- SpecSoft 3D Live (sold as hardware or software, with a free version available for download at http://www.spectsoft.com/)

**STEREOSCOPIC IMAGE ANALYSIS**

As a digital art, stereoscopy benefits from image processing applications. For live production or shot setup, new gear is arriving to the market, with a new set of tools that can be described under the umbrella name of computer-aided monitoring. These tools go far beyond the passive image processing we are used to, involving a lot of genuine science and bits of artificial intelligence in the process.

The purpose of image analysis is to understand the relationship between a left and right image: how the cameras are positioned, and how the lenses tweak the light toward the sensors. Where, in 3D, are the objects in the field of view, what do they look like, what are they made of, and how do they move? Basic
image analysis will provide correspondence points, whereas complex analysis will generate a significant amount of metadata describing the scene. The results can then be used just as monitoring helpers, or to control the rig’s actuators and eventually to correct the 3D in the image domain.

**State of the Art in Stereoscopic Image Analysis**

Stereoscopic image analysis is a complex science; it takes a PhD just to understand the jargon. And to be totally honest with you, we are still on our way there. Still, we consider it to be a central piece of the science and future of stereoscopy, with applications in every production and postproduction phase.

The purpose of image analysis is to understand, down to the subpixel level, the relationship between the left and right images. This means identifying visual features in images and pairing them perfectly. It is a binocular artificial-vision process. When enough points have been found and matched, it is possible to figure out precisely what’s wrong with your camera alignment and how to fix it—with a speed and accuracy you’ll never match. Nobody can estimate the amount of a rotation, down to subpixel accuracy, just looking at a pair of pictures for one second. The real-time correction systems presented in the next section are based on such analysis, followed by various types of image manipulation that produce perfect stereoscopic images. Even if it is a complex multidimensional mathematical resolution, we, simple artists, can describe the process as 2D. See it as a Photoshop layer manipulation, not a Maya object rendering.

There are two big challenges in the field: one is to filter out the false pairs of homologous points, which is when the computer puts together two unrelated features. You want as many couples as possible, with the best accuracy ratio, and these are antonymic expectations. How you sort them out and where you draw the line are the factors in the secret recipe that determines the value of an image analysis system. The second challenge is to run the whole process in real time, and to generate a smooth correction—which is especially important in live 3D, when one cannot delay images. At the very best, your computer will always be working on a frame that has already gone on air by the time it has figured it out.

Both domains are quite crowded when you look at the science side of it, with a respectable amount of publications on the topics. On the other hand, it’s quite a desert when you search for actually available solutions. To our knowledge, there were only two companies in the world with actual products on their shelves back in 2008, when this book was first planned: 3ality Digital in the United States and Binocle in France. Since then, three contenders have presented their solutions: the Fraunhofer Heinrich Hertz Institute (HHI), with its STAN (STereoscopic ANalyzer), codeveloped with the German production company, KUK; Sony, with its MPE-200, brought to fame by the FIFA 2010 World Cup coverage in 3D; and, in Spain, Kronomav, with its rigs and
analyzers. They are working at interfacing their computers with other vendors’ motion controls.

Applications of Stereoscopic Image Analysis

A stereoscopic image analyzer is like a new crew member with an incredibly acute visual system and a very crude understanding of what entertaining pictures are. What would you do with such a trainee? Here are a few examples of ideal tasks:

1. Providing stereographers, DPs, ACs, or directors with enhanced monitoring tools.
2. Accelerating 3D rig setup, with alignment suggestions, on screen, or via motion control.
3. Automatizing a 3D camera unit’s interaxial and convergence based on depth settings.
4. Coordinating a 3D camera fleet, with shared depth settings.
5. Automatically screening 3D content for stereoscopic defaults.

Computer-Aided Monitoring

Stereoscopic image analysis makes it easier to live with the truth that one has to see 3D to make 3D. The problem is, of course, that you may not always, if ever, have a good enough 3D display on set. Let’s talk size: even the best professional 3D displays will not give you the finesse you need to evaluate an image that’ll be seen on a 40-foot screen. Current 3D monitors come in 10-inch, 22-inch, and 46-inch sizes, on which evaluating parallax, divergence, and vertical disparities is quite a challenge. Experienced stereographer Vince Pace explains that a 46-inch 3D screen is all he needs to evaluate 3D imagery. Before you get to that level of experience and confidence, however, you will benefit from the help of dedicated 3D vision computers.

Another problem with field 3D monitoring is brightness. 3D visualization systems have an overall efficiency ranging in the 15% to 20% range, which means that with your 3D glasses on, you’ll get a picture five to six times dimmer than you’ll get on an equivalent 2D display. Here are the visual modes you’ll find on a typical 3D-Computer Aided Monitoring:

1. Show accurate parallax, vectors.
2. Identify screen plane, foreground, and background using color-coded tags.
3. Show the depth bracket of the shot on depth histograms.
4. Compare the measured parallax against the target screen size.
5. Run various 2D image quality controls in a 3D mode.

Disparity Vectors

In this mode, homologous points are shown with an arrow linking them. This mode allows for evaluating the parallax and pinpointing most geometric defaults, such as rotation and the causes of various vertical disparities, including zoom mismatch.
Disparity Tags

In this mode, homologous points are shown with false color points or icons showing the amount and direction of parallax. A cross-vendor color-code convention settled on blue tags for out-of-screen elements, green tags for the screen plane, and red tags for the infinity plane. Elements placed in the spaces between these arbitrary planes are in shades of cyan and yellow. Purple is sometimes used for flagging objects with a negative parallax stronger than the desired maximum foreground.
The threshold is computed upon dialed-in screen size, and sometimes additional factors including viewing distance and divergence limit. If you prefer, you can type in your reference positive and negative parallax values in pixels or percentage of image width.

Binocle’s Disparity Tagger displays geometric shapes that enhance visual information. Strong positive or negative parallaxes are shown with triangles pointing upwards or downwards. Screen plane is tagged with squares. Vertical misalignment is shown by rotating the shapes to align their base with the corresponding vector they represent. The size of the shapes represents the confidence of the computation. The areas of best efficiency happen to be the subject of interest, with sharp and detailed textures. In other words, your hero is more likely to be strongly tagged than the background—how convenient!

**FIGURE 3.17**

*Depth Analysis on 3ality SIP.* In this monitoring mode, many parameters are tracked, and a visual representation of the “depth budget” is shown in the lower third.

*Image courtesy of 3ality Digital.*

All vendors include a depth histogram to evaluate the depth composition of your shot. When an image must be corrected in post, it has to be rescaled. The image processor can infer the correction, the amount of resizing, and the resultant framing. That final active image can be shown on the monitor.
Comparative 2D Analysis

All the image analysis tools you used in 2D will be needed in 3D too, not only to analyze the 2D images for defaults, but sometimes to check for 3D coherency. Typically, sharpness or edge detection is useful to check the focus and iris match between cameras. Overlaid vector graph and RGB histograms are useful tools for matching the light level and color balance on cameras.
**COMPUTER-AIDED 3D CAMERA UNIT SETUP**

During the shot preparation, once you have put your cameras and lenses into the 3D rig, you want to align them, including matching the lenses and camera electronics. Intelligent disparities analysis will streamline that process.

**Optical Axis Alignment**

You can always align your cameras visually, and then check it out by getting the disparity vectors computed. What about the other way around? Run a computing pass and rely on the vector map to align your rig. This way you don’t have to guess how much rotation and translation you should correct first. You do still have to figure out which knob to turn in what direction, and how much—and remember that all of them influence the others, so you’re in for a good deal of trial and error. This is one of the reasons why stereographers have a personal relationship to their rigs. After a while, you learn to do it quickly. Soon, image analyzers will work with CAD models of rigs, cameras, and lenses, and will be able to suggest lists of alignment steps and values for the very setup you are working on.

On a motion-controlled rig, it’s even simpler. The computer will automatically align the cameras and run another image analysis until the optical axes match each other.

**Color Matching, White Balance, Aperture, and Gain**

Every camera has its very own color behavior, and even if your cameras are perfectly matched, the half-mirror is shifting them unevenly. Visually balancing whites and levels can be a harassing task; just shoot at your reference white and let the computer tell you the RGB values it gets. Alternatively, you can match the color histograms on the 3D image processor. Image analysis can do all that in a single pass by analyzing a reference chart, or a real image of the lighted set. Some software suites in development can generate color setups to be loaded in the cameras.

The aperture and gain are trickier. Asymmetric apertures will generate asymmetric focus depth. Image analysis can check out images for sharpness and match them, but the exposure levels on the sensors may not match. On a beam-splitter, it will not. From that point you have two options: add filters, or set gain.

**LENS PAIRING**

Lenses, mirrors, filters—there are many elements on the light path that are not digital, and therefore may never be perfectly matched. Quantifying the discrepancies on set will give you a chance to correct for as much as possible before the raw picture gets recorded and the stereo balancing has to be done in post. Do a favor to the FX team: get the most you can from the lens pairing. After that point, all corrections are time, money, computers, and long nights.
Z-LUT Generation

Once the rig is aligned, it’s time to take care of the lenses. Some lens vendors offer matched zooms sold by pairs. At some point, lenses may even have their LUTs computed at the factory and embedded somewhere in an on-board memory. In the meantime, you’ll be working with lenses that do not match, and you’ll have to take care of that, in shoot preparation and in postproduction. Just like color matching, you’d have a much better time in post if you did it well in prep.

Before image analyzers stepped in, Z-LUTs were generated by zooming all the way in and out on an alignment chart. Because zoom telecentry follows various sinusoidal curves, you need to realign the cameras at as many in-between points as needed, until the motion-control system tracks it closely enough. Automatic Z-LUT is a great tool that shrinks this process from 30 minutes to a handful of seconds.

Ongoing improvements in technology will soon allow you to get rid of alignments patterns. This is important for live productions, when the zooms are known to drift away and their LUTs may have to be recalibrated. Thermal dilatations are the most obvious issues, for example, following nightfall during soccer matches or heat generation by dancing crowds in a concert location.

REAL-TIME RIG CONTROL

If you have a fully motion-controlled rig and a powerful image-analysis system, the benefits of linking them are obvious. You’ll be able to compensate for the optical defaults, thermal dilatations, and even automatically set depth on your shots.

Based on Pre-Established Decision Rules

This basic function is the most popular and easiest to apply. It can be performed by semi-intelligent control units added to a basic rig, or, alternatively, can be part of a high-end rig. Binocle integrates them inside the camera socket that encloses the motion-control actuators, and 3ality has a control unit mounted on the side of the rig. Sony added remote control of rig motion to version 1.2 of its 3D Processor Box.

Based on Real-Time Computation

This advanced function is the most advanced of all. At the time of this writing, only Binocle and 3ality offer fast enough image analysis systems to interactively set the depth on a camera.

3D Rig Motion-Control Command Protocol

Until 2010, image analyzers and motion-controlled rigs were produced together in single vendor solutions. With the Sony 3D Box and the STAN, the situation is changing. Sony is using an Element Technica protocol that has been made public to this end. For the European Project 2020 3D, Jordi Alonso of MediaPro is working with Kronomav to develop a universal rig interface.
FUTURE APPLICATIONS

At the moment, the industry is only scratching the surface of the potential benefits of image analysis. With the increase in computing power and finessing of the software, more tools will soon come to fruition.

Roundness Evaluation

On a 2D TV, all balls are round. In a 3D TV program, some are spherical, flat, or elongated, all depending upon the screen size, viewing distance, depth position, and inter-optical distance. Because of the effect of the viewing distance, it is not easy to set the inter-optical and convergence solely based on a visual check on a small screen. At some point, image analysis will be able to communicate to you an evaluation of the “roundness” of a shot by identifying known objects like human faces, and comparing their 2D on-screen size with their depth range.

Depth Chart Generation

With 3D TV channels opening around the world and 3D crews shooting and editing on the same scope, at some point the contents will be shared over contribution feeds, bought at festivals, or advertised in catalogs. Chief Technical Officers will be asked for a solution to ensure the legal department that no bad 3D is reaching the audience’s eyes, and image-analysis robots will deliver the depth chart of any program that is run through them.

Breakout Circuit on Networks Operations Center

How can you make sure that no uncomfortable 3D reaches the audience? Set an image analyzer on the program feed, and make it fall back on 2D if it detects disparities that may hurt your legal standing. Have it log the average and maximum disparity at all times, and back up the data with your legal department for the day someone brings your channel to court for having damaged his 3D vision.

Surface Mirror Optical Description

We have mentioned the defaults of surface mirrors many times, and knowing exactly the default of a system brings you halfway to overcoming that default. An image analyzer coupled with adequate test charts and procedures can describe the very optical behavior of a mirror. That signature can then be injected into the VFX suite to match it.

Limits of Image Analysis

Like many computer-based magic boxes, there’s some discrepancy between the sales pitch and the actual delivery of automated image analysis. So far, artificial intelligence is still an oxymoron. There are two well-known issues in image analysis: the impossibility of finding homologous points that don’t exist, and the importance of input quality.
EDGES OF THE FRAME

On the lateral edges of the frame, we have asymmetrically occluded image areas. Whatever is seen by only one camera, in front of or behind the stereoscopic window, cannot be matched to the other point of view. This is especially unfortunate, because this is the place where you’ll have the most disturbing invasion of the viewer’s space; it is the area where people and objects enter the frame and generate 3D-impairing window violations.

You’ll want the image analyzer to detect these and fix them by reconverging the cameras or automatically setting a floating window. Unfortunately, these are precisely the points an image analyzer cannot figure out, and for a good reason: such surfaces do not have homologous areas. And that’s the very problem, of course.

The solution would be to shoot and analyze an oversized picture and frame in. There are many reasons to do that in 3D, and image analysis is just another one.

IMPERFECT INPUT, IMPERFECT OUTPUT

The system does with 3D feeds what your brain does with your retina’s neuron streams; it looks for identifiable details that repeat in each view and translates their angular distance into estimated depth. This process relies on the assumptions that the two images are equivalent in all characteristics but the point of view, and all the disparity vectors are horizontal. If this were right, it would mean you may not even need to correct your images.

The point is, because the photography of two cameras is not perfectly matched in illumination and color, feature matching is impaired. Consider a situation in which you are trying to match a surface that has a luminance gradient and one of the cameras has a higher light level. The same luminance vector (getting darker) can be associated with both photographic mismatches and depth mismatches. How to decide which is which when you are just a processing unit?

Furthermore, there’s a feedback loop here. Fixing the photographic parameter will influence the depth computation, and vice versa.

- Photographic mismatch shifts the light and color of feature points.
- To perfectly fix photography mismatch, one needs to identify homologous points.
- To perfectly identity homologous points, one needs to fix photography mismatch.

Among the source quality impairments affecting stereoscopic analysis, we’ll mention low light conditions, motion blur, genlock inaccuracy, and rolling shutter orientation.

Existing Products

There are currently four high-end products on the market, from 3ality, Binocle, Sony, and HHI/KUK. (We will discuss this more in the next section, on image correction.)
Recently, word came out that various lower-end projects built on video acquisition dongles and software image analysis are on their way to NAB 2011. It is to be expected that all video acquisition or format conversion hardware vendors (Gefen, Matrox, Blackmagic, Miranda, etc.) will come up with some sort of 3D software tool to complement their hardware offerings.

Cel-Scope3D is an example of a hardware-independent image analyzer: [http://www.cel-soft.com/celscope3d/](http://www.cel-soft.com/celscope3d/)
Hamlet Video VidScope-3D: [http://www.hamlet.co.uk/products/software/vidscope-3d/](http://www.hamlet.co.uk/products/software/vidscope-3d/)

**FIGURE 3.20**
Cel-Scope3D Analyzer.

### STEREOSCOPIC IMAGE GEOMETRY CORRECTION

The next step, when you have a robust, fast, and accurate image-analysis system, is to have it take care of fixing whatever disparities were not fixable in the camera at shoot time. We are not talking of colorimetric correction, but geometric corrections. All mechanical systems have a finite accuracy, and no rig will move a full-size digital cinema camera within a micrometer precision. This includes rotation, magnification, and optical center shift—everything that is not fixable down to the subpixel, or that can shift during the shot session or live event.

When one has a chance to see 3D correction demonstrated, it’s almost like looking through newly cleaned glasses for the first time—as if there were a blurry curtain we didn’t pay attention to, until we wiped our glasses and all of a sudden the world was crystal clear. The uncorrected image feels strangely
okay, until the correction system kicks in and you feel like someone turned the light on while you were reading in dim light.

It may seem like overkill to correct what appears as nothing more than visual background noise—but, in reality, there’s no such a thing as visual noise in 3D. What you get are retinal disparities, increasing the audience’s visual system workload. Making sense of this imperfect 3D requires physical realignment at the eye level, and additional fusion complexity in the visual cortex. This means causing visual fatigue for the fittest, and headaches for the masses—not a good selling point for your 3D content. You may think that a couple pixels of offset are not enough to generate real discomfort, but you’d be incorrect.

In addition, as soon as you mix graphics with your images, you mix perfect geometry—the CGI—with imperfect geometry—the live feed. In a post-produced feature, it means hours of labor to fine-tune it, to make the blue-screened actors match the CG backgrounds and stick on the floor. In a live 3D feed, like a sporting event, it means that the audience has to choose between looking at the scores or at the players.

Post tools will be presented in Chapter 4, and live correction tools are listed hereafter.

State of the Art in Image Correction

Stereoscopic image correction used to be called *stereo balancing*, or, when it was manually done, *stereo grooming*. It involved a little bit of vertical, a chip of rotation, two bits of zoom, and a lot of time. Its automatic counterpart consists in taking the disparity vector map from the image analysis, and figuring out a way to make all the vectors horizontal. In most cases, a 4-point transformation is enough, the very same operation you do in Photoshop when you independently ALT-move the four corners of your selection.

The challenge is to make sure that you keep the screen plane consistent, that you do not rotate the scene in any direction by skewing one view, and that you do all this smoothly. The current analyzers are not yet computing at full frame rate, but the correctors do. You have to interpolate the analysis results, based on a frame that’s already gone in the pipeline, on a handful of frames, waiting for the next batch of vectors. Some systems use the motion-control metadata to help that guessing stay on track with the camera movements.

More optical defaults like nonlinear distortions (for example, barrel, pin cushion, or moustache) are actually not significant on broadcast and digital cinema lenses. They could be fixed if needed when postproduction requires it. So far, it has made no sense to implement them in stereoscopic geometric correction tools.

Applications of Stereoscopic Image Geometry Correction

The ultimate goal of image correction is to get from the premise that “any 3D shot is an FX shot that needs to be aligned and corrected” to the premise that “any 3D camera unit produces good 3D, in accordance with a director’s depth
request.” In a comparison already mentioned, the U23D concert movie was in postproduction for over a year in 2007, and then in 2010, just a few years later, the Black Eyed Peas 3D show was a live broadcast.

This was accomplished by having a device instantly and automatically do what a stereographer would do: detect, identify, qualify, and fix the image disparities, from geometric errors to colorimetric mismatch—in short, all of them.

**EXTEND HIGH-END RIG ACCURACY**

When you just drop your cameras into your 3D rig, you get misaligned pictures by tens of pixels. Mechanically registering your cameras will bring you within a few pixels of accuracy. A good monitoring tool, patience, and experience will drop the vertical errors under three pixels.

Image analysis will allow you to drop to a pixel of accuracy, and motion control will make the process faster. That’s as long as you do not change the shot, move the rig, or follow fast action. Then your accuracy drops, and the disparities get noticeable again. There will surely be vendors claiming that their rigs are stiff enough and that their motion-control systems are strong enough, so image correction is not needed. They basically claim that the laws of physics do not apply to their products. For those who believe in gravity and inertia, and plan to broadcast what they shoot, image correction is the way to get sustained subpixel image registration.

**REPLACE MOTION CONTROL ON LOW-END RIGS**

At the 2010 NAB, Sony introduced the 3D solution they would use to produce the FIFA World Cup a few month later. They surprised the industry by selecting rigs that would not offer dynamic vertical registration, and instead would rely on their image-processing systems to fix it. The existing state of the art was to control one camera pitch to compensate for the vertical component of the zoom optical axis shift. Sony has since introduced the ability to motion control rigs in later software versions.

**LIVE CORRECTION OF DEPTH ACCIDENTS ON BROADCASTS**

Who can say nothing will ever go wrong in his live broadcasts? Maybe someone who set up a 3D image correction system on its NOC delivery chain. Set your limits, and leave it to the disparity detector to decide if it can fix the shot, rather than suffering a fallback to 2D.

**FUTURE APPLICATIONS OF IMAGE CORRECTION**

The current state of the art is only to generate an optically accurate image, as it would have been produced by a camera placed at the perfect position needed for the shot. Other renderings are possible, in the continuation of what is currently done in stereoscopic CGI animation under the name of *multi-rigging*. Two options seem feasible in the very near future: scene depth compression, and object depth extension.
Scene depth compression would deal with excessive IOD by bringing backgrounds and foregrounds toward the screen plane. This is possible because there’s no occlusion revelation. Still, the edge detection needs to be perfected.

Object depth extension would deal with insufficient IOD by inflating the volume of objects. This would require increasing the accuracy and resolution of depth maps generated from the disparity vectors.

Limits of Stereoscopic Image Geometric Correction

In our imperfect world, magic boxes can only get you so far. Whatever light that did not reach the sensors will be quite hard to synthesize on a computer. Furthermore, image geometry correction inherits some of its limits from the image analysis it relies on, as well as from later interaction issues with post-processing tools.

GEOMETRIC CORRECTION IS NOT VIEWPOINT SYNTHESIS

Image correction takes care of two potential camera position errors: the origin and destination point of the optical axis. Changing the aiming point of a camera is quite easy; just shift the image, as we do when we reconverge stereoscopic footage. Changing the sensor position is another story. That’s what we do when we interpolate viewpoints in postproduction.

Image correction can only correct what the camera is aiming at, not where it is. You still need to get to perfect camera elevation. Horizontal misplacement will generate inappropriate amounts of IOD and convergence angle. That’s wrong, but it is not as much of an issue as vertical misplacement generating unresolvable retinal disparities. Claiming to be able to fix vertical shift is just as bold as claiming to fix incorrect IOD.

GEOMETRIC CORRECTION AFFECTS VISUAL EFFECTS WORKFLOW

There’s no sensible sharpness loss in image correction; bicubic interpolation does not really show up once you have applied a couple of video compressions along the distribution channel.

The issues come with match-moving software that may not be able to understand images whose optical center is not in the middle of the frame. Pixel Farm’s 3D tracker gets its stereoscopic fame partly from the fact that it can work with corrected images. Otherwise, the process is to record raw images with image-analysis results, or corrected images with image-correction parameters. In any case, always flag images that have been tweaked at shoot time.

Existing Gear

3ALITY

The first to market was 3ality Digital, with its SIP 2100 and the upper-scale, multicamera unit, the SIP 2900. It offers extensive monitoring modes, DVI and SDI connectivity, frame-compatible format conversions, and Web remote
control. It also drives the rig motors and corrects geometries. It does all this using two PowerPCs and four Xilinx Spartans, for a total processing power of a teraflop. It uses mostly 2D image analysis to control the cameras and make sure the best image is produced from the start, while 3D geometries are computed and corrected in non-Euclidian spaces.

This piece of equipment has been under development for many years and is used on 3D live productions like BSkyB Sports, and on movie sets. Even though it was designed to work with 3ality rigs, it is often used with other systems, especially on feature productions.

The 3ality approach is to put the quality burden on the rigs and cameras, and to minimize the amount of corrections to be applied afterwards. These quality expectations are to correct "mechanical shifts" that will always occur during the production of a live show. When you keep zooming in and out, and pan the camera side to side for hours, you apply momentum on cameras, rigs, and lenses. Experience shows that your mechanical registration will not last until the end of the show. To that end, the image has to be analyzed and corrections have to be sent to the rig.

**BINOCLE**

Binocle is a French company that has been doing stereoscopic research and productions for more than ten years. They have teamed with the French National Institute for Research in Computer Science and Control (INRIA), known for creating the RealViz FX suite. At NAB 2009, Binocle introduced its image processing system, the Disparity Tagger. It is a regular Windows computer with SDI in and out capabilities. At NAB 2010, the first generation of real-time geometric correction was presented. Binocle’s philosophy is to embed the rig’s motion control and computing resources inside an all-in-one socket that holds the camera. As a result, the image analyzing system is totally independent and does not interact with rig automation.

**KUK**

The STAN is the result of collaboration between the Fraunhofer Heinrich Hertz Institute (HHI) and KUK productions. It was presented at NAB 2009 and has since been adapted for P+S Technik rigs. It is now available.

The STAN approach is to use information coming from the rig’s motion control to help with the time extrapolation part of the geometric correction. It’s also a software solution that runs on a PC platform. It has a simplistic but intuitive user interface.

**SONY**

Sony’s product is composed of a multipurpose digital production system, the MPE-200, and a 3D acquisition specialized software, the MPES-3D01. The hardware is based on the Cell BE2 engine, with four HD inputs and outputs, and the software is developed in the United States.
Sony’s approach is to assume that all stereoscopic defaults can be corrected in the image domain, and that therefore a powerful enough image processor can replace a complex rig. The current version of 3D software does not use any image analysis or motion control; all the geometric and colorimetric corrections have to be entered manually during the system setup. The MPE communicates with the lenses via the camera’s HD/SDI interface to receive focal length and focus distance information.

Sony has announced a new version that will be able to communicate with rigs using Element Technica’s command protocol.

**FIGURE 3.21**
Camera Alignment on Sony MPE-200 with MPES-3D01.

**KRONOMAV AND MEDIAPRO**
Kronomav presented its 3D rigs at NAB and CES and developed the first image analyzer, which is already for sale. The Spanish company has now teamed up with the TV producer MediaPro to create a full 3D TV production platform. This includes a new version of the analyzer that will be introduced at NAB 2011, which works both at controlling the rig motion and correcting the pictures in the image domain.

The project manager for MediaPro, Jordi Alonso, is currently working at developing a standardized communication protocol between rigs and image analyzers. Doremi, StereoTec, and Element Technica have shown interest in the ongoing preliminary work. The hardware implementation is on CUDA processors on a regular PC.