NIST Special Publication 500-XXX

Imaging Criteria and Test Methods for Qualification of Iris Cameras

FIRST PUBLIC DRAFT

National Institute of
Standards and Technology
U. S. Department of Commerce

Homeland Security
Science and Technology

Dan Potter
Patrick Grother
Elham Tabassi

I N F O R M A T I O N
T E C H N O L O G Y

Information Technology Laboratory
National Institute of Standards and Technology
Gaithersburg, MD, 20899-8940

May 8, 2013

U. S. Department of Commerce
Rebecca M. Blank, Acting Secretary

National Institute of Standards and Technology
Patrick Gallagher, Under Secretary for Standards and Technology and Director

Table of Contents

Reports on Information Technology

The Information Technology Laboratory (ITL) at the National Institute of Standards and Technology (NIST) stimulates U.S. economic growth and industrial competitiveness through technical leadership and collaborative research in critical infrastructure technology, including tests, test methods, reference data, and forward-looking standards, to advance the development and productive use of information technology. To overcome barriers to usability, scalability, interoperability, and security in information systems and networks, ITL programs focus on a broad range of networking, security, and advanced information technologies, as well as the mathematical, statistical, and computational sciences. Special Publication 500-series reports on ITL's research in tests and test methods for information technology, and its collaborative activities with industry, government and academic organizations.

Disclaimer

Certain commercial entities, equipment or materials may be identified in this standard to assign field numbers to registered vendors or to describe a procedure or concept adequately. Such identification is not intended to imply recommendations or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

Acknowledgments

The Iris Device Qualification Test protocols have been developed through the financial sponsorship of the Department of Homeland Security, Science and Technology Directorate (DHS S&T) Homeland Security Advanced Research Projects Agency (HSARPA) under the technical stewardship Mr. Arun Vemury, Program Manager. The contents and opinions expressed within this document do not necessarily reflect the views or policies of the Department of Homeland Security or any of the individual components that make up DHS, nor does the mention of trade names, commercial products or consultants constitute endorsement or recommendation by these agencies.

1 Introduction

This document defines an iris camera evaluation and qualification test funded and developed by the Department of Homeland Security, Science and Technology Directorate (DHS S&T), Homeland Security Advanced Research Projects Agency (HSARPA), and NIST. These efforts are intended to aid in the consideration of iris biometrics for applications ranging from building access control to expedited and secure processing of foreign nationals at, air, sea, and land ports of entry. Device qualification is necessary for mitigation of risks associated with capture of poor iris images. Its role is prior to technology deployment, and it may be used in cost benefit analyses, and facilitate an efficient procurement process.

The document is part of a larger tiered process that attempts to separate the lower level device performance characteristics from how these potential capabilities work in real applications where human behavior may significantly influence image quality as it relates to iris biometrics (Figure 1). This document presents a series of optical tests which are specifically designed to evaluate the device-level image quality aspects of iris biometric capture device. It does not attempt to address the more complicated yet important human interactions between a device, an attached system, the human subject, the device operator, and the operational environment. These aspects are amenable to test under, for example, the ISO/IEC 19795 Biometric Performance Testing and Reporting standard. This first tier of the test, the Iris Device Qualification Test (IDQT) provides content for down selection decisions of devices prior to human-in-the-loop testing for US Government applications.

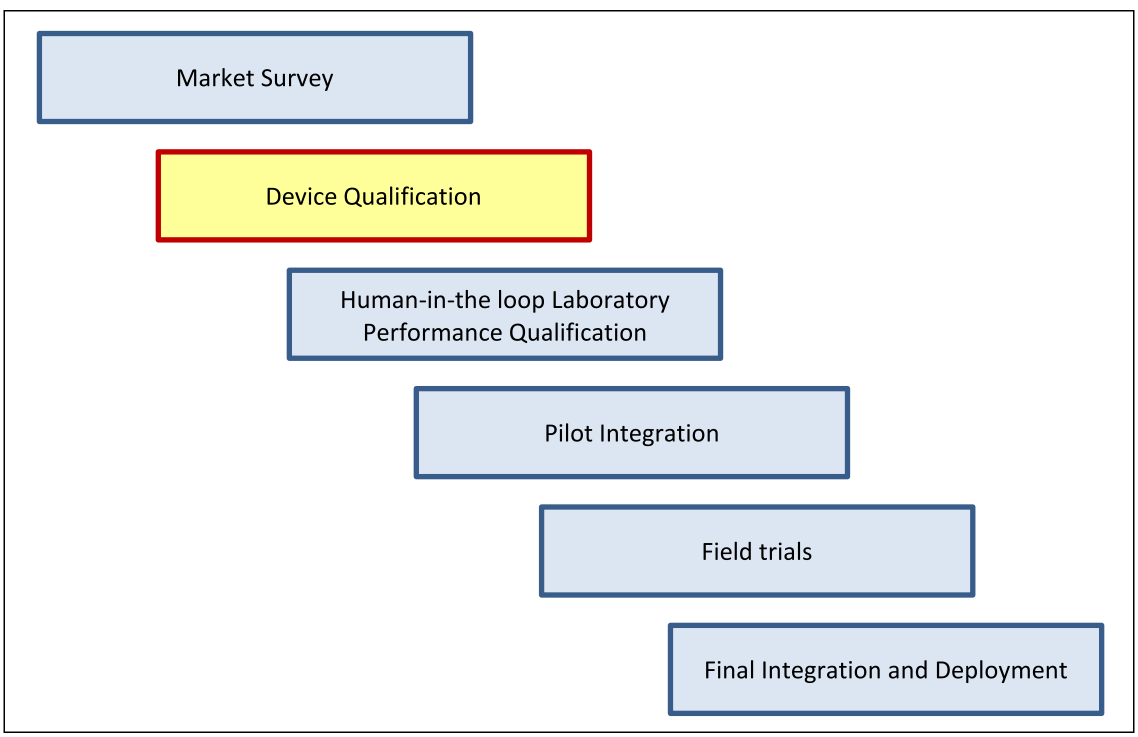


Figure 1 The procurement process for iris biometric technology for an arbitrary application follows a serial tier from an initial market survey of possible device solutions to a full deployment and integration into existing infrastructure. This document focuses on the test and procedures needed to fulfill device qualification.

The motivation to separate the device qualification as a prerequisite for further testing is both technical and economical. Human performance testing in a laboratory or a field location is often more expensive and time consuming compared to device level testing, and is not repeatable due to population variance effects. Creating a device qualification pre-requisite prevents expending resources on testing devices whose peak imaging performance falls short of application requirements. On the technical motivation, the specific products of the device-level evaluation give insight into the strengths and weaknesses of a given device removed from human elements, allowing comparative studies of devices to be used in later stage testing and evaluations. It also allows for a confirmation of manufacturer's performance claims relative to application-specific requirements and/or the requirements and recommendations stated in ISO standards[1].

The following sections present the IDQT plan. Section 2 outlines the scope of the test, and the specific metrics used in the test. Section 3 reviews the testing equipment required to conduct the test, and the rationale behind the design of the test equipment. Section 4 reviews the specific procedures involved in each test in collecting, and processing the image and diagnostic data, which are used to formulate the suite of quality metrics. Finally, Section 5 outlines the criteria for qualification, along with their technical justification.

2 Test Overview and Scope

The information gathered from the IDQT is, by design, intended to measure “peak” imaging capability using metrics relevant to the signals contained in near-infrared images of the human iris used by recognition algorithms. The test is repeatable and should therefore be able to discern whether or not a device, removed from all human factors, is suitable as an iris capture device.

The IDQT follows a conventional image quality assessment procedure, whereby a series of calibrated optical patterns are presented to the imaging device undergoing evaluation followed by the computation of relevant metrics from the captured images. With the output metrics from this test in hand, comparison can be made to qualification criteria to make a judgment on the suitability of a given device for further consideration. The test has been designed to accommodate measurements which exceed the current guidelines with the rational that the information content requirements of future matching algorithms may increase over time. However, as expanded upon in Section 5, the exact qualification criteria used in a given project would likely depend on the application.

Where possible this document specifies the use of image quality assessment hardware and software that is standardized, published and widely available. However, in many cases, existing techniques were found to interfere with the operational image capture processes and this necessitated the development of equipment and techniques dedicated to iris device qualification. Other non-imaging

[1] These are:

- The published iris interchange standard ISO/IEC 19794-6:2011 Information technology -- Biometric data interchange formats -- Part 6: Iris image data which revises and replaces the 2005 iris standard.
- The draft iris image quality standard ISO/IEC 29794-6 Information technology -- Biometric sample quality -- Part 6: Iris image which defines acceptable properties of iris images.

tests are included, such as the characterization of the on-board illumination in the context of eye safety standards, and the basic physical measurements of dimension and weight. The optical properties of the eye and face region are incorporated into the test targets to ensure the evaluation devices capture images of test targets as they would of real irises.

The collection and analysis procedure is designed to accommodate different capture modes which may influence how a particular device is evaluated. The intent here is to support qualification devices running in different capture modes. Capture-mode diversity is distinct from how a device can be used in different applications, and although this test covers different capture modes the qualification criteria are application dependent. For example, one device could be used for collection of images for national scale deduplication, and for outdoor access control. The qualification criteria could be different for the two applications.

2.1 Device Capture Mode Classifications

Specific considerations of the iris image capture process have been made in the IDQT collection and analysis procedure to accommodate a variety of possible operating modes. While conventional enrollment devices produce single images of the left and right eyes, a new generation of iris capture devices explicitly tradeoff capture speed for accuracy to facilitate applications such as high-throughput access control. In such cases, the number and quality of iris images output by a device for one presentation of a subject may vary and it is not necessary that all output images are of a high quality. Such “opportunistic” modes of operation continuously feed images through an operational iris recognition algorithm essentially using a (fast) iris matching process as a quality filter. Aspects of this mode of operation have been incorporated into the test procedure in an effort to make the IDQT a minimally biased evaluation applicable to a wide variety of applications.

A definition of a capture event is used which encompasses a range of operational modes. The start of a capture event is defined according to manufacturer’s documentation, typically initiated by a command through a supplied demonstration application, or from the appearance of a face in a specified capture volume. The end of a capture is defined according to the manufacturer’s documentation, marked by a message received from the device by the control computer for example. If the end point is not well defined by continuous capture devices for example, a period of ten seconds is allowed for the capture process to take place for each capture. This definition is applied to the following classes of capture modes used in the test outlined below.

Single Image Mode

The single image mode of operation results in the device providing one iris image for each presentation. Each image taken during the collection process in single image evaluations is used in the evaluation analysis.

Opportunistic Stream Mode

For each presentation, multiple images are output from devices in opportunistic mode. Out of the series, the image with the highest performance metric is used in the evaluation analysis.

2.2 Liveness Detection

Some devices may have liveness detection features which may prevent the release of an image, or compromise the capture process. In order to test such devices, the manufacture shall disable these capabilities.

2.3 Acceptable Image Formats

The IDQT accommodates a wide variety of output image types, not restricted to the widely used 640x480 eight-bit greyscale format compatible with commercial matching algorithms. This test can accommodate any format which can be translated into the numerical pixel values of the images and their correct orientation relative to the focal plane. Image formats include ISO/IEC 19794-6, PNG, JPEG 2000, TIFF, and bitmap or any format for which access to the relevant data elements is supported and defined. If problems are encountered reading the supplied image format, or if the format contains multiple channels when one is expected, the device manufacturer shall afford a remedy after appropriate notification. The format of the output images shall be included in the report. Although some basic information regarding the format of the images output from iris capture devices, this test does not replace a formal standards-conformance test to ISO/IEC 19794-6.

It should be noted that although the IDQT does not test explicitly evaluate device output format, this fact does not imply that DHS or other Federal applications will not have output format requirements at later-stages in the down selection process.

2.4 Interoperability

Currently there are many available commercial matching algorithms and iris capture devices. This availability has increased substantially in recent years. Interoperability of the images taken from different device manufactures is important in federated multi-party applications. Interoperability facilitates competition and minimizes long term risk associated with depending on one proprietary, non-interoperable commercial solution.

While device-level testing and qualification is intended and expected to improve interoperability, there is as yet limited understanding of root-causes of cross-device recognition accuracy degradation. Particularly because human-interaction effects may play a role, tests such as those defined in ISO/IEC 19795-4 Performance Interoperability Testing are most suited to quantifying cross-device recognition accuracy.

The metrics collected in this document's qualification processes will support analysis of cross-device interoperability effects. In particular, by correlating results with empirical recognition trials, it might reveal device level root-causes of non-interoperable behavior. This motivates the repeatable, purely optical test of imaging capability given in this document.

Although this test may give the US Government insight into the possible root cause of potential interoperability shortfalls, there are no explicit reporting mechanisms related to interoperability resulting from the IDQT.

2.5 Levels of qualification

Enrollment versus Verification Criteria

Typically, application requirements dictate that the database of enrollment images is of a higher quality relative to the image used for verification. It is conceivable that this device qualification may have separate criteria for enrollment and verification applications. However, the definition of an “enrollment quality” image is application specific, in particular on the false match and non-match performance levels the application requires. For example, an application involving just 1-1 verification may require less quality compared to a program which is attempting to de-duplicate a nation's individuals. Also, the determination of enrollment quality lies heavily on human presentational issues, such as occlusion, eye gaze, and pupil dilation state which are not possible to control on a device level test. Therefore, although the output metrics from this test support qualification of a device (for enrollment, say), they do not ensure that all images will be suitable for enrollment. The qualification criteria from the IDQT may provide a basis for in a down selection process of an enrollment device versus a verification device, but because these criteria are application dependent the IDQT does not fix a definition for an enrollment quality image versus a verification quality image.

2.6 In-Scope Measurements

The following subsections review the specific device level elements that are in scope for this test.

Spatial Frequency Response

Spatial Frequency Response (SFR), analogous to the Modulation Transfer Function (MTF), is a metric used to quantify the amplitude attenuation and phase change of imaged sine wave patterns of varying frequency through an optical system. Quantitatively, at a given spatial frequency the SFR is the modulation of a sine wave at that frequency seen in the image divided by the modulation in the original object plane. The MTF can also be expressed as the complex amplitude of the Fourier transform of the Point Spread Function (PSF). The Contrast Transfer Function (CTF) is a similar metric to the MTF of an optical system, but is calculated on the basis of a square wave pattern rather than a sine wave pattern. The SFR is relevant to iris biometrics as feature extraction algorithms use a limited range of spatial frequencies from the iris pattern seen in near infrared (NIR) images. Each algorithm may use a specific set of spatial frequencies, Removed from the spatial frequency response of iris matching algorithms, criteria for the SFR metric can be arrived at based on ISO standard recommendations of the spatial sampling rate, Nyquist sampling theory, and assumptions regarding the point spread function as outlined in appendix xxx.

Pixel Scale

Pixel scale describes the number of pixels occupied by a physical distance in the object plane, in this case in the physical plane of the iris. The pixel scale is important as it relates to the spatial sampling rate which determines whether the PSF is adequately sampled. Irises typically range between 10.2

and 13.0 mm with an average of about 11.8 mm [2]. ISO recommendations based on results from the IREX III study[3] suggest that images which should have a pixel scale with at least 160 pixels across the iris diameter giving a pixel scale between 15.7 and 12.3 pixels/mm. A refined guidance for pixel scale is outlined in section 5.

Iris Albedo Texture SNR

The information used in iris biometrics is comprised by subtle light and dark pattern variations intrinsic to the iris, outlined further in Section 3. Physically, the signal source is the NIR albedo variations across the iris that has amplitudes of less than a few percent on top of a fairly dark albedo with median of around 15%. The metric developed in section 4 specifically created for this test measures the effective response of the filtering effects of a typical iris encoding and matching process using iris-like texture in an attempt to provide a near "bottom line" measure of how well an optical system records iris-like features.

Field Distortion

The field distortion can be defined as the position dependence of the pixel scale in the image plane. There are no quantified ISO recommendations for this metric. It should be noted that some forms of field distortion may be mitigated through the iris segmentation and pseudo-polar transform operation carried out in common matching algorithm paradigms. The metric is based on an integrated area of distortion relative to three nominal length scales used in the three levels of qualification criteria outlined in section 5.

Greyscale Gain Linearity

The greyscale gain linearity refers to the relationship between the number of NIR photons recorded in a pixel and the expressed digital value of the pixel. This relationship may be non-linear depending on the camera processing and/or sensor characteristics. Although there are no ISO standard recommendations regarding gain variations, non-linearity may be the source of interoperability issues, and segmentation failures for some algorithms expecting a linear signal, for example.

This quality aspect is measured using a supplied target which has regions of known NIR albedo covering the range of interest for iris biometrics.

Object Plane Albedo Resolution

Iris information is conveyed through spatial variation in greyscale. Insufficient sampling of the albedo distribution will lead to a reduced number of grey levels in the captured image. The object plane albedo resolution refers to the smallest change in albedo that can be detected by the device. The metric is based on the number of grey levels spanning the 0.8 to 0.25% range of NIR albedo targets. Because this magnitude of the albedo resolution depends on the scale of the feature size, the metric uses the three scales used which define the three levels of qualification from section 5 (0.75mm,0.375mm, 0.25mm) to define area sizes used in the metric creation.

---

[2] Andre Caroline, Effect of Corneal Diameter on Soft Lens Fitting, *Contact Lens Spectrum*, Vol. 17, No. 4, 2002.

[3] P. Grother, G.W. Quinn, J. Matey, M. Ngan, W. Salamon, G. Fiumara, C. Watson, Iris Exchange III, Performance of Iris Identification Algorithms, NIST Interagency Report 7836, April 9, 2012. http://iris.nist.gov/irex

Ambient Light Mitigation, supplied illumination corneal reflection map

The cornea reflects about 2-3% of the incident ambient light. If this reflection overlaps with the iris from the camera’s viewing angle, the reflected scene may be superposed with the iris information. This test evaluates a given device's ability to mitigate ambient light noise, and maps out the reflection scene from a device's supplied illumination.

Exposure Time Estimation

A test is included which estimates the device's effective exposure time using a series of spatially separated lights which are turning on and off with a known synchronization. Exposure time has some relevance to iris biometrics as shorter exposure times translate into a higher acceptable error in tracking subject motion. This metric should be taken with the caveat that some systems may have sophisticated closed or open loop tracking systems which result in low tracking errors even with high subject velocity and/or acceleration. This test does not explicitly evaluate device performance as a function of subject motion. However, such a test could be accomplished through placing the test targets on motion platforms which cover a realistic range of subject and/or operator motion.

Supplied Illumination Spectrum

Depending on the eye color, the human iris exhibits different patterns as a function of wavelength. Blue eyes do not have a front facing pigment layer, so the characteristic pattern is similar from the visible to NIR wavelengths with features becoming lower in contrast with higher viewing wavelength. Brown eyes however have a pigment layer which is essentially opaque to visible wavelengths, but optically thin to NIR wavelengths. In the NIR, the pattern underneath this visibly opaque layer contains higher contrast and higher albedo features. The appearance of the iris pattern changes with wavelength, particularly brown eyes are dark and relatively featureless in visible wavelengths so iris imaging almost always uses NIR wavelengths between 700nm and 900nm. Additionally, the limbus boundary has lower contrast at 900nm and this can affect segmentation.

The output metric reports a low spectral resolution (20nm) measurement of the spectral irradiance between 700 and 900nm integrated over sufficient number of capture events. This test assumes that the main source of iris illumination is from the device itself.

The draft ISO/IEC 29794-6 standard has three requirements on spectral composition in the range [680,920]nm. These are:

1) That 90% of power on [680,920] resides in [700,900];
2) At least 35% of power in [700,900] resides in [700,800]nm and
3) At least 35% of power in [700,900] resides in [800,900]nm.

Editor's NOTE: This document does not currently formalize a requirement on spectral composition for iris capture devices because no good study exists. Comments are sought on whether and how to formulate requirements in this area.

Eye Safety

This test includes the spectral characterization of the supplied illumination, as well as irradiance measurements to compare with the ACGIH eye safety threshold limit values[4].

Capture Volume

The capture volume is the physical space which an iris capture device can produce an image which satisfies a qualification criteria. In the case of iris biometrics, some devices will not instigate or release the results from an image capture sequence without a subject present in a specified capture volume. The capture volume will be measured relative to the stated volume from the manufacturer's specification.

2.7 Out of Scope

Although some of the information gathered in this test can reveal important performance details of a given device, the test laboratory shall not report backwards engineering on specific device parameters. Namely device characteristics such as the aperture size and F number, sensor quantum efficiency, sensor noise characteristics, image stabilization techniques, or post processing methods used are not explicitly evaluated.

The test does not explore the influence of operational environments including temperature variations beyond nominal room temperature, humidity variations, and long term effects of outdoor exposure to dusty and corrosive conditions.

The IDQT does not specify tests for ruggedization, durability, and vulnerability to malicious exploitation.

3 Test Design and Equipment

The IDQT test procedure and equipment are designed with the goal of accurately measuring the peak imaging capabilities in the context of iris biometrics in a way which minimizes biases that would misrepresent the performance of one device over another. An example of a possible source of bias is that some devices use the optical features of the face and eye in the capture process to automate tracking and focus and a test using optical targets, which does not accurately include these important features at a level of realism required by the iris camera, may interfere with the capture process for these devices. Also, the iris texture relevant for iris biometrics is confined to a fairly narrow range of NIR albedo variations as discussed previously. Most off-the-shelf imaging targets consist of high albedo contrast patterns which do not represent real iris features. Besides the possible capture bias mentioned above, diagnostic measurements using COTS targets may overestimate the performance in recording information available from the iris. In short, without targets that accurately represent real face and iris features, the test may misrepresent performance of certain devices relative to others.

[4] ICNIRP Statement on Light-Emitting Diodes, Implications for Hazard Assessment http://www.icnirp.de/documents/led.pdf

The following section reviews the details test equipment developed by DHS S&T/HSARPA specifically for the IDQT and the rationale behind the test design.

3.1 Human Face Optical Target Mount

A three dimensional representation of a human face is required for this test for two reasons. First, a number of iris capture devices use the features of the face to locate and track the eyes for iris imaging. Without a realistic presentation of at least generalized face features, some devices will fail to perform image capture. Secondly, light scatters from the nose and eye socket, which subsequently reflects from the corneal surface and back into the camera aperture. These nose and eye socket reflections may constitute a noise source, and therefore a model representing the 3D profiles of the face features, along with at least the general scattering features of the skin is required to test these aspects of iris biometrics.

These face features need to be reproducible and well defined to enable the repeatability of the test. After an extensive exploration of the near infrared scattering properties of various readily available materials, the decision was made to use 3D printing technologies to create the face and optical target mounts. These models are shown in Figures 2 and 3.

The model of the face was chosen using software and a face dataset from a commercial source. The face used in the IDQT model was generated from an average face representation from a large and diverse population of faces. This 3D model was scaled preserving aspect ratio such that the interpupillary distance matched the a nominal mean of 63 mm[5]

Figure 2 A 3D representation of an "average" human face with a mounting solution for model eyes which have specialized optical targets (Figure 3). The face model has a ¼-20 thread on the bottom to facilitate mounting to optical bench hardware to accurately control stand-off distance and presentation angle to evaluation devices

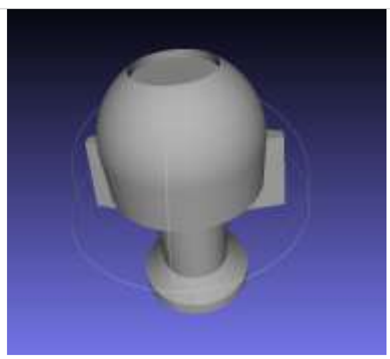


Figure 3 An design for the eye target blank which fits inside the face mount shown in Figure 2. The piezography printed pattern and cornea simulation solutions fits inside the front facing indentation in the design. The reflective properties of the 3D printed model is similar to that of the human sclara as seen in near-infrared wavelengths.

[5] Dodgson, N., "Variation and Extrema of Human Interpupillary Distance" Proc. SPIE Vol. 5291, Stereoscopic Displays and Virtual Reality Systems XI, A. J. Woods, J. O. Merritt, S. A. Benton and M. T. Bolas (eds.), 19–22 January 2004, San Jose, California, USA, ISSN 0277–786X

The face is augmented with sets of artificial eyebrows and lips with a different contrast relative to the skin regions to provide realistic signal for face detection algorithms.

3.2 Optical Target Design and manufacture

The optical targets used are mounted in the eyeball-like base model of Figure 3. This was manufactured from a digital model using 3D printing technology. The front of the base has a 12 mm diameter circular indentation to accommodate the various optical target patterns which are chosen to be the size of the typical human iris[2]. The model of the optical eye target has two keyed tabs which fit into grooves in the side of the holding cylinders on the back side of the face mount. This can accommodate 90 degree axial rotations of the eye target to control against possible systematic influences of the target manufacture. One of the two keys is colored and the groove positions are numbered 1-4 to consistently position the targets, and keep track of the rotation orientations used.

3.3 Review of Specific Diagnostic Image Target Patterns

The patterns are printed on paper using piezography carbon-based ink using a high resolution inkjet printer calibrated to produce controlled NIR albedo values as a function of spatial frequency. The target patterns used in the IDQT, shown in Figure 4, have been designed to incorporate and represent the features used in iris biometrics, namely the NIR albedo contrasts of the iris pattern, and the pupil and limbus borders. Explanations for each target used in the test follows in subsections below.

A number of target patterns used for image analysis have been created to fit in a circular area with a diameter of 12.0mm, or inside the nominal area of a typical iris. These patterns fit inside the eye model described in section 3.2 and have been designed to work with analysis algorithms designed to extract image quality information relevant to iris biometrics. Before use the albedo of the printed patterns are validated using calibrated albedo targets and a large format NIR camera system which produces images with a modulation of greater than 0.95 at a frequency of 20 lp/mm and below, or significantly higher frequencies than conventionally used in iris biometrics.

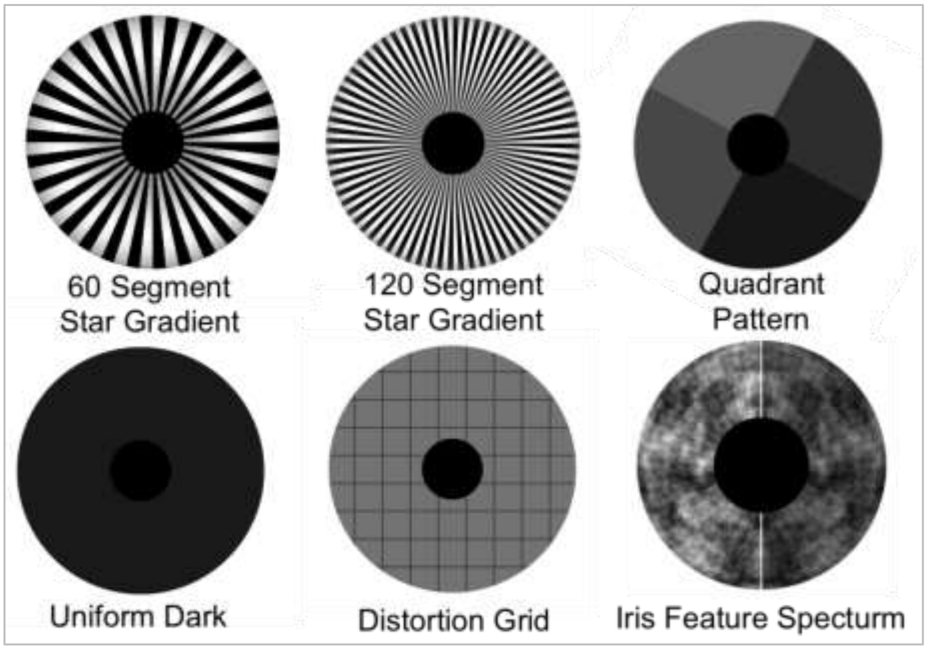


Figure 4 The six target patterns used in the IDQT are shown above. Each target has a diameter of 12.0mm and a pupil diameter of 4 mm.

3.3.1 Quadrant Pattern

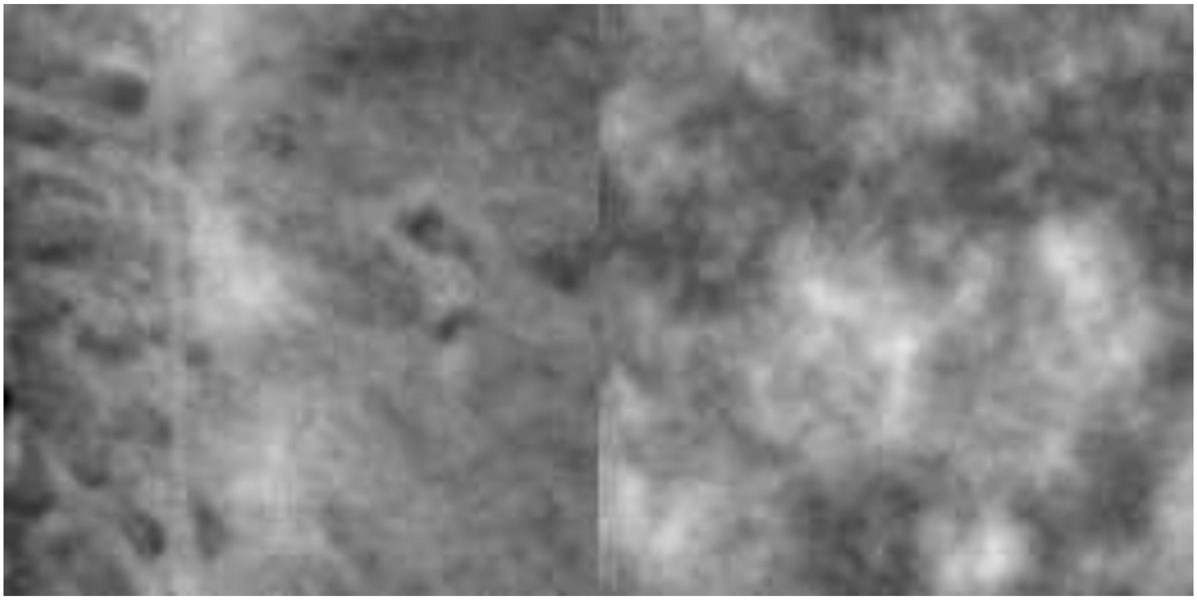
The quadrant pattern consists of four even-area 90^o slice sections each assigned a different calibrated NIR albedo value. The values used in the four regions (0.08, 0.14, 0.20, 0.26) are chosen to span the range of albedo found in the iris. The edges have a rise with a half-width at half-maximum on the order of 10 microns to facilitate slanted-edge MTF measurements using the line spread function in the range at frequencies lower than ~10lp/mm. The four areas are large enough to produce independent noise measurements as a function of spatial scale spanning scales from 0.2mm-2mm, with the target paper texture limiting noise measurements at higher spatial frequencies. The resulting albedo variations are characterized and accounted for in the statement of the metric for linearity and slanted edge MTF resulting from the test.

3.3.2 Iris-Like Feature Spectrum

This target in intended to represents the near infrared albedo features of the iris. The goal of this target is to produce a pattern with an albedo amplitude distribution which is in the range of real irises. To guide the pattern design, characterization of the near-IR iris pattern was conducted using a portion of the University of Bath iris image database[6], and using nominal values of the image scale using a nominal value of iris diameter of 11.8 mm. Iris texture not obscured by eyelids, lashes, or bright specular reflection was Fourier analyzed in ~2.5x2.5 mm square bins as seen in Figure 5a. Although individual bins may exhibit departures from the average, a fit to the radially averaged power spectra (Figure 5b) of the bins fit the profile of a random Gaussian field with an amplitude

[6] University of Bath Iris Database (http://www.smartsensors.co.uk/information/bath-iris-image-database/)

distribution following a power law exponent value of -11/3 with feature size within a range between 0.07mm to 1 mm, as shown in Figure 6.



Example NIR Iris Texture **Random Gaussian Field with PLC=-11/3**

Figure 5 An example of the 2D NIR iris texture shown in comparison with a random Gaussian Field with a power law coefficient (PLC) of -11/3. Across many iris images sampled at frequencies with high MTF (greater than ~80%) the value of -11/3 was found to be the best general fit to the measured iris texture power spectra.

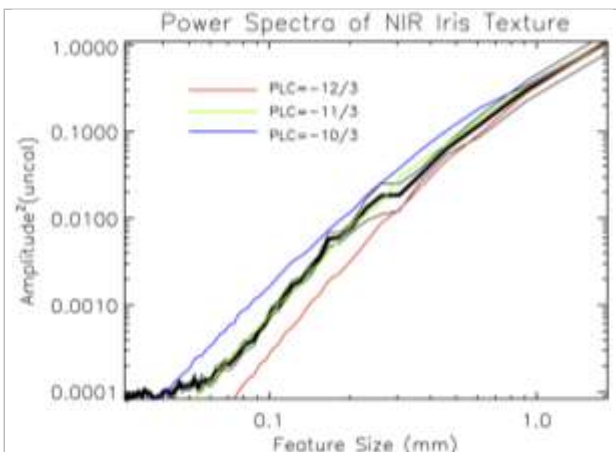


Figure 6 An example power spectra of the 2D iris texture information as recorded by near infrared imaging shows (black line) that the iris texture follows a power law distribution with a coefficient (PLC) equal to -11/3. The different color lines shows different PLC values for comparison, with the blue line showing a slightly more even distribution with PLC=-10/3, and the red line a steeper distribution with PCL=-12/3. (need to show the plot representative of

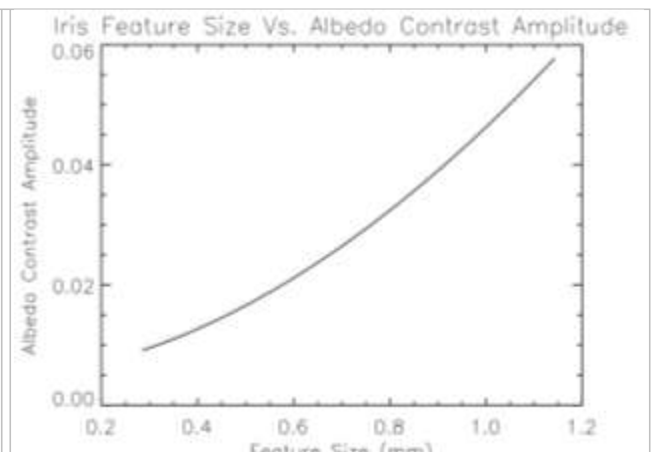


Figure 7 The profile of the albedo contrast as a function of feature size is shown from the resulting power $r^{-11/3}$ relationship established in Figure5b and matching the absolute. (consider plotting 1 sigma or some nominal value above the average… the low contrast features will be masked out via fragile bit filters)

3.3.3 Gradient Contrast Star Patterns (60 and 120 segments)

These targets are designed for direct measurement of the CTF using contrast targets with albedo characteristics representative of those measured for the human iris. The 60 segment gradient contrast star pattern covers frequencies ranging from 0.8 lp/mm to 3.5 lp/mm. The 120 segment pattern covers a higher range between 6.5 lp/mm and 1.6 lp/mm. There is a slight overlap in coverage between the two targets to confirm results with separate physical targets for the level II and level III criterion. Following the -11/3 power law relationship of the power spectrum of iris texture, the albedo contrast amplitude representative of typical iris features as a function of feature size is shown in Figure 7. The dark and bright regions are set to have a difference of 0.03, or the typical value of iris features which are around 0.7mm in size. Although this strategy may overstate performance on typical iris texture at smaller sizes, setting uniform albedo values with radius for a given angle simplifies both the CTF extraction algorithm, and creating the target. Also, although the average albedo contrast of iris texture at high spatial frequencies may be lower than the 0.03 value of the test target, the iris features observed at scales at the lower range of our consideration ~0.1mm are at or above this contrast.

3.3.4 Uniform Dark

A uniform, dark background is ideal for mapping out the distribution of reflected light from the cornea overlapping with the iris area. The uniform dark target consists of a uniform albedo of 0.08 across the whole of the iris region. The texture of the paper results in a departure from uniformity at small spatial scales. The residual RMS albedo error from the printed paper target under uniform illumination is taken into consideration in the formation of metrics.

3.3.5 Distortion Square Grid

The square grid target consists of lines of width 0.1mm spaced 1.3mm apart. The line pattern is printed with an albedo of 0.05, with the background at 0.16. This pattern is used to estimate the field distortion of image capture devices. The field distortion is expressed as absolute spatial displacement error relative to a linear model presented normal to the optical axis. The output metric is expressed both in object plane Cartesian and the conventional “pseudo-polar” coordinates used in iris biometrics.

3.3.6 Exposure Time Target

This target is comprised of 14 optical fibers implanted in the iris area evenly spaced in two radial layers. The fiber diameters are 0.5mm, and are coupled to the light output of 14 separately controlled LEDs which are optically isolated from each other with rise times on the order of 5 ns. The brightness of the LEDs are small compared to the typical return from an iris capture device, and can be controlled if the brightness interferes with the capture process for devices which use the corneal reflection. The 14 lights sequentially turn on and off, configured to be at 5ms intervals. In a given image of this target, the number of fibers illuminated is an indicator of the exposure time up to about 60ms.

3.4 Representation of Corneal Reflection

Two different methods are used to represent the reflection characteristics of the cornea. The corneal reflection from device-supplied illumination can be used as feedback for tracking and focus control so it is important to replicate these physical characteristics in the target to avoid bias. In addition, the ambient light reflection of the corneal surface may be a significant source of noise in iris images, and so this needs to be incorporated into the test as well. One set of IDQT optical targets use a smoothly curved surface of UV curing glue placed in the pupil center with a chosen index of refraction and radius of curvature to simulate a corneal reflection. A second set of IDQT targets use magnesium fluoride (MgF2) coated spherical lenses with a radius of curvature of 7.75 mm, which is similar to that of the human cornea, tear reflection surface. Versions of each optical target are made using both techniques and incorporated into the data collection procedure.

All patterns are used with the UV cured glue corneas as these give an unobstructed view of the target pattern. The MgF2 lens targets are only used with the uniform, dark and iris texture patterns for suitability assessment for different ambient light environments.

Neither target design accurately represents the "red-eye" return aspect of the human eye observed when the illumination sources are near the entrance aperture of the imaging device. Thus, the target design cannot accommodate devices which may use the red-eye return aspect of human eyes in the capture process.

3.5 Vibration Environment

Mechanical vibration is controlled through a floating optical bench with an active vibration dampening system, which provides a vibration environment less than VC-C velocity based environmental vibration criterion which is below the operational theater ISO standard.[7] Suitability for different vibration environments are not covered in the IDQT. Considering the wide variety of vibration environments possible including a vibration environment suitability test would not be practical. Potential issues regarding vibration would presumably be discovered at later stage field testing.

3.6 Eye safety and Device illumination Characterization

The IDQT measures the evaluation device irradiance with a calibrated irradiance meter at the range of standoff distance for recommended use from the vendor, with measurements taken between this distance and the closest standoff distance a subject could conceivably come to the device illuminator. For addition illumination characterization, one eye target has a 1 KHz response photo-diode incorporated into the pupil region, as well as an eye target with a fiber collimator which feeds into a NIR spectrometer. These devices allow for a both a wavelength and time dependent pulse characterization which will be included in IDQT evaluation reports. The pulse characterization is meant to provide additional input to augment exposure time estimates, to encompass the possibility that some capture devices may use pulsed illumination as the effective method to "freeze" motion

[7] Gordon, C.G., *Generic Vibration Criteria for Vibration-Sensitive Equipment*, SPIE,3786,22, 1999

rather than exposure time. In addition, eye safety threshold limit values as outlined in the ACGIH TLVs and BEIs [8]Wavelength characterization

3.7 Ambient light control

Iris cameras may require some degree of ambient light to detect a subject and initiate a capture process. However, there are illumination scenarios which may inhibit image capture and recognition. The index of refraction difference between the cornea surface and the air causes a small fraction of the light incident on the cornea to be reflected. Depending on the relative angle of the incoming light and the orientation of the iris and the camera, these reflections may interfere with the iris pattern recorded in an image.

To both accommodate the requirements for device operation, and to evaluate how devices may mitigate unwanted light sources, an ambient light control system has been designed to simulate the influence of a range of ambient lighting environments. This system consists of a number of diffuse and compact broadband sources as well as calibrated near infrared sources which can be positioned at different distances and angles relative to the target location.

The IDQT includes simulation of three ambient lighting environments to match possible applications Table 1. These are indoor without sunlight through glass, indoor with sunlight through glass, and outdoor.

Table 1

Ambient Light Scenario	Lux Reading (Human Response)	NIR Irradiance (700-900nm) mW/cm^2
Indoor, no Sunlight through glass	50-500	~1.e-3
Indoor, sunlight through glass (same as outdoor in shade)	2500-5000	~1.e-2
Outdoor (consider outdoor shade+ outdoor)	25000-50000	~0.1

3.8 Communication with Device Manufacturer

Sources of possible bias have been identified and mitigated in this test plan to accommodate iris capture devices. However, it is possible that important operational details of a given device may preclude a device from being tested, and that the manufacturer may not be able to reveal mechanisms to augment the test without revealing commercially sensitive information to the lab. In light of this possibility vendors are contacted prior to evaluation, and given this document to review to identify any areas of possible bias in the test procedure. If a device fails to produce an iris image for a given IDQT test target, the details of the failure may be given to the vendor to give an

[8] TLV and BEIs Based on the Documentation of the Threshold Limit Values for Chemical Substances and Physical Agents and Biological Exposure Indices, ACGIH 2013 (http://www.acgih.org/)

opportunity to provide a possible solution or information which may provide a cause for alternative means of passing into scenario testing.

3.8.1 Information useful from the manufacturer

The following is a list of information expected from the device manufacture that will ensure that the test represents the manufacture's intended operational use.

1. Set-up and installation instructions
2. Operating instructions
3. Directions on how to access out images and quality information if available
4. Hardware Requirements
5. Stand-off distance from corneal surface or other facial reference point
6. Operational capture volume
7. Ambient light requirements
8. Location/definition of optical axis
9. Power requirements
10. Other operational requirements

Any other relevant information from the device manufacturer will be taken into consideration to facilitate the testing process.

4 Test Procedures

4.1 Overview

This section outlines the procedures for carrying out the IDQT. All but one metric involved in the test is derived from the output image of the collection device, with the exception being the illumination characterization. For the image based tests, after the diagnostic images are collected, they are analyzed by image processing algorithms to produce output metrics which are then compared to a set of qualification criteria outlined in section 5.0. The information is reported in a format that includes at least sample images used in producing the metric.

4.2 Pre-test preparation

For each device being evaluated by the IDQT, an IDQT initiation form shall be submitted as a pre-requisite to begin testing. This form is intended to be the mechanism for the device vendor to convey information needed to successfully carry out the test, and to assert what the category and operational mode or modes best describe the device. A point of contact from the vendor will be designated to establish a technical dialog with lab facility in the event of possible technical problems that may be encountered while conducting the test. A non-disclosure agreement (NDA) between the testing facility and the device vendor is also initiated at the pre-test preparation stage.

4.3 Data Collection

The core of the IDQT procedure involves the collection of images from the six categories of test targets. Before target collection begins, a check is performed to make sure that the evaluation device is working according to the manufacturer's directions, and that the output images are accessible for analysis. The collection procedure is not the same for all targets, with some targets requiring more angles of rotation than others, or for example the use of the ambient light control shield. For each image target collected, a corresponding algorithm, or set of algorithms is applied to produce the metrics used in the criteria comparison. Table 2 illustrates the relationships between the targets and the corresponding output metrics.

Each test collection consists of 5 collections taken with 4 target orientations. A metric value can be computed from single captures, so statistics can be performed through the 20 collections to estimate the error of the test metrics, with possible systematics detected from the changes in rotation orientation of the target.

Table 2

#	Image Quality Metric	Required targets
1	Spatial Frequency Response	60 Segment Star Pattern
		120 Segment Star Pattern
		Quadrant Pattern
2	Pixel Scale	All targets
3	Greyscale linearity	Quadrant Pattern
4	Greyscale Resolution	Quadrant Pattern
		Low contrast Star Gradient
		Feature Spectrum
5	Iris Feature SNR	Low Contrast Star Gradient
		Feature Spectrum
6	Ambient Scene Corneal Reflection Noise	Uniform Dark
		Feature Spectrum
7	Instrument Illumination Corneal Reflection Noise	Uniform Dark
		Feature Spectrum
8	Exposure Time	Exposure Time Target
9	Illumination Eye Safety	Photo-Transistor Target
10	Illumination Wavelength Characterization	Fiber Spectrometer Target

4.3.1 Pre-Collection Setup

Procedure for device and target placement

The face model is designed with a standard ¼ - 20 mount which accommodates standard tripod mount hardware and COTS optical mount hardware. Using this mount, the face target is secured to a vibration-controlled optical table at one end. The iris camera undergoing evaluation is mounted relative to the optical targets according to the manufacturer's directions for stand-off distance. An alignment laser mounted in the center of an eye model is used to align the target system to the optical axis of the iris capture camera. If the optical axis of the iris camera is not well defined or specified by the manufacturer, the best estimate is chosen by the test operator, using feedback from the initial try of the test target images.

Note for devices too large to fit on the optical table

Some iris capture devices are too large to fit on the 4'x8' optical table or are designed with the notion that the subject will be walking through the capture volume. If a device cannot be placed on the optical table, an alternative mounting solution will be used according to the device form factor and mount feasibility. In some cases it is expected that the device will be placed on the floor of the laboratory room, or using extensions to the optical table. These cases will not fully benefit from the vibration control of the optical table and in the case of unexpected performance failures, root cause analysis may be performed to make sure vibration is not the cause. In the case of a walk through application, improvised solutions will be employed with feedback from the vendor.

Procedure for excluding "failed" Captures

The following are criteria which would invalidate an image for consideration in the test.

1. A capture (illumination flash and appropriate feedback) which results in no output images from a device.
2. An output image which is not visibly recognized as a target image
3. An image which satisfies Chauvenet's criterion to exclude a data point using the initial set of 20 images (per target) to develop baseline statistics.

In line with the goal of measuring "Peak" imaging performance, if there are a number of failures for a given test target, attempts to recapture images will be made until a set of at least 16 images are obtained with no defined outliers.

4.3.2 Detailed Image Collection Procedures

The following procedures are to be carried out after the device and face mount have been positioned, and all manufacturers' recommendation for operation have been confirmed. Lighting for the collection uses a commercially available fluorescent light evenly distributed behind the capture device with a nominal value of 100 lux, in line with the "indoor, windowless" ambient light category but without significant texture in the scene presented to the target. NIR Irradiance, lux from a conventional photography light meter, and spectrometer measurements from ~600-1200nm are taken before data collection for each device to ensure the ambient lighting environment does not significantly change from device-to-device.

For each test, and regardless of the mode of the test, the same type of target will be used in both eyes simultaneously, with both targets being rotated together in each rotation angle iteration for devices that capture both eyes from one presentation.

Collection Procedure for Image Quality Targets

The following steps are to be followed in the collection of the images for targets including the 60 and 120 segment star gradient contrast pattern, the quadrant patches target, the iris-like feature spectrum target, and the square grid target. The order that the different target patterns are taken does not matter.

For two-eye capture devices:

1. Place the cornea free version targets in each eye slot at rotation position 1
2. Perform 5 capture attempts
3. If successful, repeat 5 capture attempts for each of the other 3 rotations, if unsuccessful reattempt up to 3 times for each failed capture attempt until 5 successful captures are collected

For one-eye capture devices:

Carry out steps 1-3 above for each eye separately.

After analysis, and elimination of outliers, if less than 16 images remain in the qualification set then recapture attempts (up to 3) will be made until at least 16 images per target are obtained, with at least 4 per rotation angle.

Collection Procedure for Ambient Light Qualification

The ambient light test requires the patterned scene illumination hardware be attached to the optical table, and the illuminated patterned scene consisting of contrast patterns which, projected from the corneal surface give features between 0.2 and 2 mm as measured at the object plane of the iris target (cornea surface). These patterns may interfere with iris texture. The ambient light scene is designed to range in illuminance as measured from the target with 4 fixed levels (20, 100, 5000, and 50,000 lux). The image collection procedure is as follows:

1. Place the illuminance meter eye module in the face target, and adjust the ambient light to one of the three light levels being tested.
2. With the corneal surface versions of the uniform dark eye targets in place in each eye slot, follow steps 1-3 in the previous section, capturing 5 images for each of the 4 rotation angles for each eye.
3. With the corneal surface versions of the Iris texture target in place in each eye slot, follow steps 1-3 in the previous section, capturing 5 images for each of the 4 rotation angles for each eye.

Collection Procedure for Exposure Time testing

Images of the exposure time target with the blinking lights powered on are used to estimate exposure time. The target is validated with controlled imaging using a both a long exposure to ensure that all lights are operational, and with known short exposures to confirm that the blink duration for each light is 5ms. Five images from each eye need to be captured for the analysis to average out effects from devices which may use progressive scan rather than global shutter modes.

Collection Procedure for illumination Characterization and Eye Safety

The illumination characteristics will be measured using a calibrated irradiance meter, in conjunction with a spectrometer with optical apertures mounted in the face target near the iris region. The time signature of the illumination sequence during the target imaging will be recorded and compared to eye safety guidelines[8].

4.4 Application of Image Processing Algorithms

The result of a successful image capture process for a device in a single image per capture attempt mode is a total of 530 images. Considering set-up time and confirmation of the data quality, the entire data collection takes about 4-6 hours. Once collected the images are processed through a series of diagnostic algorithms which produce the IDQT metrics. The following subsections provide a high level review of the output metrics.

4.4.1 Pixel Scale

The pixel scale refers to the spatial sampling rate of pixels expressed in the object plane coordinates, or referenced to the scale of the iris. The units are in number of pixels/mm. This sampling rate is important as it sets limits in the realm of conventional imaging as to what spatial frequencies can be measured without aliasing signal. The pixel scale can be measured in a straightforward manner from any image collected in the IDQT collection sequence by using the scale of the outer diameter of the iris targets. These are automatically calculated through a segmentation algorithm and manually checked. With the exception of the capture volume testing, the pixel scale should not vary with a fixed standoff distance and this principle is used as a quality check for processing.

4.4.2 Spatial Frequency Response

In the ISO standardized notion of iris biometrics, identity is performed through a signal extraction and matching of near infrared iris images. All the information of relevance to iris biometrics is contained in the spatially varying albedo of the iris, which is recorded as greyscale image values. The ability of an imaging device to measure spatially varying signals of different sizes is referred to as the spatial frequency response. There are a number of standardized ways of measuring the spatial frequency response from an image with a known target. The most common metrics used are the modulation transfer function (MTF) using either a point source, slanted edge, or sinusoidally varying patterns and the contrast transfer function (CTF) using targets with alternating bands of light and dark.

The IDQT measures MTF using two different methods. One uses the slanted edge method (ISO 12233) using the edges at the borders of the quadrant pattern target. The other method uses the two star target patterns to formulate a direct measurement of the CTF, which is closely related to the MTF. An analysis program is run on the images and directly calculates a CTF value at each of the 60 patterns around the circle with 100 radial samples which provide the variation in spatial frequencies to form the CTF. Low order aberrations such as astigmatism and coma can be inferred from changes in the CTF values as a function of angle.

The output of this test are CTF curves at 12 angles to provide error estimates, as well as MTF values from the slanted edge tests on the four quadrant borders. Tabulated values at 1, 2, and 3 lp/mm are stated for use in the three-tiered qualification scheme.

4.4.3 Iris signal-to-Noise

The purpose of this test is to measure the effective iris feature signal-to-noise ratio. Measurements of the characteristic iris texture from the generalized definition of an iris feature shown in Figures 5-7 are used to form this metric with a generalized notion of how the iris encoding and matching process works. Namely the images of the iris targets are run through segmentation and encoding algorithms that generate iris templates. Binary templates are created using different encoding filters across a spectrum of frequencies covering feature sizes from 0.25mm to 1 mm. The filter definitions mentioned in publications are included in the test, and include the log-Gabor wavelets, DCT, and Haar encoders and include a mask defined by where the recorded signal strength is below a threshold. For each of the 20 images collected from each eye, the template is compared to a "pristine" reference template and the XOR results for each defined bit are recorded at the rotation minimum based on a conventional global Hamming Distance score formed from all non-masked bits. The output from the analysis is a series of 20 XOR result arrays based on a global minimum score, the commensurate 20 Hamming distance scores, and the percentage of the 20 values of the total area which passes the mask filter. This information is the basis for the qualification criteria outlined in section 5.

In this case where there are no differential distortions or occlusions caused by iris dilation, eye gaze, or eyelid occlusion, the genuine group Hamming distances as a function of spatial frequency are directly related to the ability of a the evaluation device to record relevant iris-like contrast features under a definition of signal in line with that used by iris biometrics. Because the target represents real iris texture, the test can evolve in time incorporating new definitions of features and matching.

Preliminary testing has shown that different filter definitions give similar results as a function of frequency, and this test may allow reduction to one filter type.

4.4.4 Field Distortion

In order to measure the field distortion the method is adapted from the ISO 9039-2008 standard *Optics and photonics -- Quality evaluation of optical systems -- Determination of distortion.* This method measures the error of the extra-axial image points compared to the nominal, undistorted image points under the assumption that the target is not tilted relative to the sensor plane. A distortion map is output from the analysis in object plane Cartesian coordinates, and in the pseudo-polar coordinates after segmentation. These maps are compared to the scale sizes corresponding to the three tiers of qualification, showing the fraction of area which distortion is greater than the feature size.

4.4.5 Greyscale Linearity, Illumination Uniformity

The IDQT uses the Quadrant Step target to make a measurement of the greyscale linearity over the relevant range of albedo representing the iris texture. The output metric from this test is the deviation from linearity to an albedo reference (greyscale values with errors versus albedo value) and the greyscale-to-albedo gain value representing the best fit to a linear model. In addition, the systematic variations with position in the eye are derived from the four rotation orientations, as well as the greyscale uniformity over each quadrant.

4.4.6 Albedo Sampling Rate

The quadrant step target, and the uniform dark target are used in a measurement of the albedo sampling rate as a function of spatial scale. This is essentially based on variance measurements taken over different region sizes of the uniform albedo regions in the targets, and a calibration from greyscale values to albedo values considering the targets have a known albedo. The sampling rate at the three scale sizes relevant to the qualification criteria at the absolute albedo of 0.12 are tabulated for the report.

4.4.7 Ambient Light Mitigation, Corneal Reflection Noise

The surface of the cornea has a different index of refraction compared to the air which results (via the Fresnel equations) in a mirror-like specular reflection. In addition to the index of refraction difference, the magnitude of the reflection off the corneal surface depends on the angle of reflection. Assuming the index of refraction of the corneal surface is 1.376, this surface reflects about 2.6 percent of the incident light for most angles of concern for views of the iris along the nominal optical axis of the eye. A point source illuminator or one with relatively small angular size results in a small specular reflection spot. Sources of light with a larger angular footprint, such as an un-obscured window to daylight seen in a dark room, result in larger projected area on the cornea.

Unwanted noise for iris biometrics occurs when the scene projected onto the cornea is of a similar or slightly greater intensity of the iris pattern. The relative brightness of the iris to the corneal reflection depends on the angular distribution of the ambient light incident on the cornea. There are certain device design characteristics which can mitigate the effects of corneal reflection noise, such as the use of narrow band optical filters matched to supplied illumination, and the illumination intensity.

The images of the uniform dark target with the corneal surface with the different levels of ambient light are used to develop a map of the scene reflected by the cornea. The maps are converted to estimated equivalent albedo, or expressed in greyscale space if it is suspected that the images undergo adaptive regional histogram equalization which would complicate the conversion to albedo. The lowest level at 20 Lux ambient is considered the “instrument only” signature map, followed by the 3 different ambient light qualification levels. These maps are compared to the nominal definition of iris signal as a function of spatial scale as one input into the ambient light qualification.

The images collected of the iris texture target with the corneal surface are compared both back to a “perfect” reference template as well as the instrument only template. The metrics are sourced from the same output as outlined in the iris signal-to-noise section above.

4.4.8 Capture Volume Estimation

The iris-like feature target is used for this exercise. Capture attempts are made throughout the capture volume as claimed by the device manufacture. If important for a given application, the full suite of image collection and analysis can be carried out. For the nominal test, captures attempts at the boundaries of the capture volume are carried out with the iris texture target, with real time feedback matching to a reference template. Discrepancies are from manufactures claims are noted in the IDQT report.

5 Guidelines for Qualification Criteria

The qualification criteria outlined below are chosen so if a device passes all outlined criteria for a given application, it would be appropriate to proceed to human-in-the-loop scenario testing. When possible, qualification criteria are formed with a foundation of algorithm performance based on studies measuring the interaction between the various aspects of device-level image quality and matching performance, such as those outlined in the IREX II (IQCE) study. However, as emphasized throughout this document, the performance of an iris biometric system can be influenced by a large number of factors beyond those specific to the peak imaging performance that is measured from the IDQT procedure. This fact complicates the construction of criteria that directly correlate with the matching accuracy from a real world collection. Commercial algorithms are developed in a competitive environment and collectively explore a range of specific techniques and so may have varying sensitivities to specific aspects of device-level image quality. It is possible that future developments in matching algorithms may use information from the iris that is not currently utilized. To accommodate flexibility in a developing industry, the qualification criteria presented here are designed to be as agnostic to algorithm performance as possible, keeping in mind a realistic notion of the information that algorithms may use in the identification process, namely the near-infrared image contrast of the iris texture, and the borders between the iris and the pupil and sclera.

Another complication for qualification is that different applications will inevitably place different performance requirements on biometric devices. For example, the requirements for a national ID program, which involves large population de-duplication (N-N), will have different definitions for acceptable performance relative to verification application (1-1). A device intended for enrollment may have a different requirement for peak imaging performance compared to a device intended exclusively for verification. The following section presents a set of qualification criteria based on the metrics resulting from running the IDQT collection procedure and analysis designed to address these challenges.

5.1 Three Tier Qualification

The IDQT qualifies devices for eye safety, and that they accurately record information required for iris biometrics with an assessment for suitability considering different ambient lighting environments. Considering the wide spectrum of possible applications, and to allow for the IDQT to maintain relevance as the industry develops, a tiered qualification system is employed.

Three qualification criteria are defined to judge the ability of a device to record iris texture in three spatial frequency intervals, with the base level (Level 1) being defined by the upper response frequency for widely used commercial iris matching algorithms available at the time of writing this document. The three levels are:

- Level I: Devices must be able to deliver images that result in a measured MTF with a modulation of more than 50% at 1 lp/mm using the IDQT targets. In addition, the encoded 0.75mm features from the bare iris feature spectrum target must be matched with a HD score of 0.1 or less to the pristine reference template for at least 95% of the collected images (1 out of 20 can fail). Each Hamming distance is only valid if more than 90% of the iris area passes the signal quality mask relative to the reference template mask. Out of the three encoder types used, the type resulting in the best score is the one used in the evaluation.
- Level II: Devices must be able to deliver images that result in a measured MTF with a modulation of more than 50% at 2 lp/mm using the IDQT targets. In addition, the encoded 0.375mm features from the bare iris feature spectrum target must be matched with a HD score of 0.1 or less to the pristine reference template for at least 95% of the collected images (1 out of 20 can fail). Each Hamming distance is only valid if more than 90% of the iris area passes the signal quality mask relative to the reference template mask. Out of the three encoder types used, the type resulting in the best score is the one used in the evaluation.
- Level III: This specification is included to support future algorithms capable of exploiting higher spatial frequency information. Devices must be able to deliver images that result in a measured MTF with a modulation of more than 50% at 3 lp/mm using the IDQT targets. In addition, the encoded 0.25mm features from the bare iris feature spectrum target must be matched with a HD score of 0.1 or less to the pristine reference template for at least 95% of the collected images (1 out of 20 can fail). Each Hamming distance is only valid if more than 90% of the iris area passes the signal quality mask relative to the reference template mask. Out of the three encoder types used, the type resulting in the best score is the one used in the evaluation.

To note the tests for other metrics such as pixel scale, greyscale linearity, albedo sampling rate, and field distortion are not explicitly used in a final qualification decision. However various combinations of these aspects of image quality may significantly influence biometric performance. Instead of making a complicated interdependent set of criteria from individual components, the IDQT is based on criteria which attempt to capture the combined influences of all the potentially significant aspects of iris image quality. In the case of a failed qualification, the individually reported metrics can be used to give information back to vendors to guide them in improving their devices to meet the desired qualification criteria.

5.2 Ambient Light Environment Qualification

In addition to the 3 level qualifications, the ambient light tests are used to qualify devices in three different lighting environments. These are: 1) indoor office without windows, 2) indoor with windows, and 3) outdoor direct sunlight (see Table 1). The fraction of cornea reflection noise from the device illumination, either from the primary specular reflection or the secondary reflections from the nose, is tallied separate from that of the controlled artificial ambient light source. A device can be qualified for each of the three ambient light criteria at each of the spatial frequency-based level tests. Qualification is passed if over 80% of the iris texture information is able to produce a match and below the albedo noise criterion to a reference template for each of the 20 images collected in each set relative to the "instrument only" set, with the "instrument only" set needing no less than 90% of the area matchable relative to the noise free reference template and below the albedo noise criterion for each spatial frequency under evaluation. For example, a level III, outdoor ambient qualification would need to accurately record the encoded 0.25mm features over at least 80% iris area relative to a pristine template.

5.3 Application Relevant Device Classification

A wide variety of applications are possible for iris biometrics. Formal qualification for particular programs would be specified externally by referring to this document. The IDQT uses the device classifications that are asserted by the vendor upon application for device qualification. If there are any discrepancies between the asserted device category and the experience using the device during evaluation, this will be noted in the evaluation report. The following categories are defined as:

1. Mobile Non-Contact
2. Mobile Contact
3. Stop and Go at a Distance
4. Walk Through
5. User Position to Fixed Focus

In addition to the 5 application specific categories, additional device categories are noted based on the output. For example, two eye capture versus single eye capture, are not specifically tested, but are recorded as a part of the IDQT reporting process based on vendors claims.

5.4 Root cause analysis for Elemental Metrics

The qualification criteria are based on the "bottom line" iris texture SNR metric, as well as the MTF metric, however a larger number of other lower level metrics are produced from IDQT testing. These are used to assess the relative margin strengths and weaknesses of the evaluated devices. The DHS S&T/HSARPA reserves the right to distribute the IDQT root cause analysis back to the vendor on a case-by-case basis to provide information useful as feedback for how they may improve their product relative to meeting the IDQT criteria.

Appendix A Example Report Format

Device ID: 123456
Evaluation Dates: SEPT. 14-15, 2013
Asserted Device Category: Mobile Non-Contact, Two eye capture, Single image per capture per eye

Criterion	Measured Metrics	Result
Level I, II, III	**MTF:**	

<table>
<tr><td>Ambient Indoor No Windows</td><td>Mod @ 3lp/mm = 0.25
Mod @ 2lp/mm = 0.51
Mod @ 1lp/mm = 0.92

IRIS TEXTURE SCORES:
<HD> level I = 0.04 with 95.4% unmasked
<HD> level II = 0.06 with 95.3% unmasked
<HD> level III = 0.11 with 96.8% unmasked

AMBIENT LIGHTING NOISE SCORES:
Level I - 98.4% to Ref., 99.0% to Inst. Only
Level II - 95.3% to Ref., 98.3% to Inst. Only
Level III – 89.1% to Ref. 92.0% to Inst. Only</td><td>Qualification at:
Level II - Ambient Indoor No Windows</td></tr>
<tr><td>Level I, II, III
Ambient Indoor Windows</td><td>MTF:
Mod @ 3lp/mm = 0.25
Mod @ 2lp/mm = 0.51
Mod @ 1lp/mm = 0.92

IRIS TEXTURE SCORES:
<HD> level I = 0.05 with 95.2% unmasked
<HD> level II = 0.09 with 94.5% unmasked
<HD> level III = 0.23 with 92.1% unmasked

AMBIENT LIGHTING NOISE SCORES:
Level I - 97.2% to Ref., 98.6% to Inst. Only
Level II – 90.3% to Ref., 94.3% to Inst. Only
Level III – 75.2% to Ref. 82% to Inst. Only</td><td>Qualification at :
Level II - Ambient Indoor Windows</td></tr>
<tr><td>Level I, II, III
Ambient Outdoor</td><td>MTF:
Mod @ 3lp/mm = 0.25
Mod @ 2lp/mm = 0.51
Mod @ 1lp/mm = 0.92

IRIS TEXTURE SCORES:
<HD> level I = 0.11 with 94.2% unmasked
<HD> level II = 0.15 with 95.2% unmasked
<HD> level III = 0.31 with 96.1% unmasked

AMBIENT LIGHTING NOISE SCORES:
Level I - 83.2% to Ref., 86.6% to Inst. Only
Level II – 74.3% to Ref., 78.3% to Inst. Only
Level III – 63.7% to Ref. 72% to Inst. Only</td><td>No Qualifications for Ambient Outdoor</td></tr>
<tr><td colspan="3">Eye Safety Check:
Continuous Illumination during exposure (no pulse)
Single peak wavelength @ 811nm, FWHM of 30nm,
Irradiance on target of 0.12 mW/cm^2
EYE SAFE</td></tr>
<tr><td colspan="3">Nominal Standoff Distance:
16.4 cm</td></tr>
<tr><td colspan="3">Exposure Time Range:</td></tr>
</table>

20ms +/- 5ms
Notes: No complications encountered for device set-up Well documented operation instructions with supplied demonstration application No capture failures with IDQT targets

Raw Image data tar file:

IDQT_SEPT152013_Device_12345.tar