<u>***Best Practice* and *Face Pose Value* Documents**</u>

Best Practice Application Level 30
(Informative)

BEST PRACTICE RECOMMENDATION

FOR THE CAPTURE OF MUGSHOTS

Version 2.0

September 23, 1997

The original version of the *Best Practice Recommendation* was initiated at the Mugshot and Facial Image Workshop that was held in Gaithersburg, MD on October 23-25, 1995. Developed as a recommendation, the implementation of the practices and principles described in that document makes the conversion of existing and ongoing photographic collections more uniform. It contains a suggested set of procedures and equipment specifications for organizations considering the purchase of new systems or the upgrade of current systems. The recommendation is not designed to render current and legacy mugshot collections unacceptable. Rather, it is intended as a means of establishing or improving interoperability between mugshot systems.

The information contained in this updated revision of the ABest Practice Recommendation@, Version 2.0, does not alter any of the individual points that were consensually agreed upon and included in the original version of this recommendation. It does provide additional details and clarifications for many of those points and has been supplemented with information regarding depth-of-field and exposure considerations.

This recommendation reflects a minimum set of common denominators. The provisions of this recommendation are keyed to the quality aspects associated with the unaltered captured mugshot image. For new mugshot images being captured, the specifications contained in this recommendation are equally applicable to real-time electronic capture of mugshots as well as the electronic conversion of photographic images. For conversion of legacy files of photographs, most of the provisions of this recommendation are also still applicable. In the future, it should be possible to add additional specifications without contradicting any of the current contents of the recommendation.

*   **POSE**

The full-face or frontal pose is the most commonly used pose in photo lineups and shall always be captured. This pose is in addition to profiles or intermediate angled poses captured to acquire perspective and other information. For subjects who normally wear eyeglasses, a frontal mugshot image should be captured of the subject without glasses. This is required due to the glare from external illumination. An additional image can optionally be captured of the subject wearing eyeglasses.

*   DEPTH OF FIELD

The subject's captured facial image shall always be in focus from the nose to the ears. Although this may result in the background behind the subject being out of focus, it is not a problem. For optimum quality of the captured mugshot, the f-stop of the lens should be set at two f-stops below the maximum aperture opening when possible.

*   CENTERING

The facial image being captured (full-face pose) shall be positioned to satisfy all of the following conditions:

-    The approximate horizontal mid-points of the mouth and of the bridge of the nose shall lie on an imaginary vertical straight line positioned at the horizontal center of the image.

-    An imaginary horizontal line through the center of the subject's eyes shall be located at approximately the 55% point of the vertical distance up from the bottom edge of the captured image.

-    The width of the subject's head shall occupy approximately 50% of the width of the captured image. This width shall be the horizontal distance between the mid-points of two imaginary vertical lines. Each imaginary line shall be drawn between the upper and lower lobes of each ear and shall be positioned where the external ear connects to the head.

*   LIGHTING

Subject illumination shall be accomplished using a minimum of three (3) point balanced illumination. Appropriate diffusion techniques shall also be employed and lights positioned to minimize shadows, and to eliminate hot spots on the facial image. These hot spots usually appear on reflective areas such as cheeks and foreheads. Proper lighting shall contribute to the uniformity of illumination of the background described in the exposure requirement.

*   BACKGROUND

The subject whose image is being captured shall be positioned in front of a background which is 18% gray with a plain smooth flat surface. A Kodak or other neutral gray card or densitometer shall be used to verify this 18% gray reflectance requirement.

*   EXPOSURE

The exposure shall be keyed to the background. Several areas of the recorded 18% gray background shall be used to verify the proper exposure. The averages of the 8-bit Red, Green, and Blue (RGB) components within each area shall be calculated. Each of the RGB means shall fall between 105 and 125 with a standard deviation of plus or minus 10. Furthermore, for every area examined, the maximum difference between the means of any two of the RGB components shall not exceed 10.

*   ASPECT RATIO

The Width:Height aspect ratio of the captured image shall be 1:1.25.

*   MINIMUM NUMBER OF PIXELS

The minimum number of pixels in an electronic digital image shall be 480 pixels in the horizontal direction by 600 pixels in the vertical direction. It should be noted that the image quality of the captured mugshots and facial images will be improved as the number of pixels in both directions are increased. However, as

images are captured with an increased number of pixels, the 1:1.25 (Width:Height) aspect ratio will be maintained.

Two considerations must be noted regarding this aspect of the recommendation. First, the normal orientation of many available cameras is the landscape format which specifies a greater number of pixels in the horizontal than in the vertical direction. Unless these cameras capture at least 600 pixels in the vertical direction, it may be necessary to rotate the camera 90 degrees. Second, the 480x600 capture format exceeds the VGA display format of 640x480. Therefore, at a minimum, an SVGA specification of 800x600 pixels will be required to display the facial image. The image will occupy less than the total number of available horizontal pixels.

* **COLOR SPACE**

Captured electronic color facial images are required. Digital images shall be represented as 24-bit RGB pixels. For every pixel, eight (8) bits will be used to represent each of the Red, Green, and Blue components. The RGB color space is the basis for other color spaces including the Y, C_b, C_r and YUV. Additional color management techniques are available from the International Color Consortium. Information regarding these techniques can be downloaded from the following URL: http://www.color.org.

* **PIXEL ASPECT RATIO**

Digital cameras and scanners used to capture facial images shall use square pixels with a pixel aspect ratio of 1:1.

* **COMPRESSION ALGORITHM**

The algorithm used to compress mugshot and facial images shall conform to the JPEG Sequential Baseline mode of operation as described in the specification approved by the ANSI X3L3 Standards committee. The target size for a JPEG compressed color mugshot image file shall be 25,000 to 45,000 bytes.

*   FILE FORMAT

The JPEG File Interchange Format (JFIF) shall contain the JPEG compressed image data. The JFIF file shall then be part of the transaction file for interchange that conforms to the requirements as contained in ANSI/NIST-CSL 1-1993 and ANSI/NIST-ITL 1a-1997.

Best Practice Image Capture for SAP Levels 40, 50, and 51
(Informative)

Introduction

This set of "enhanced best practice recommendation" (EBPR) clauses is a set of constraints. These constraints can be categorized into four types of requirements: scene, photographic, digital, and format. Scene requirements refer to the content, subject and background in the image. Photographic requirements refer to lighting, focus and other constraints required for photo capture. Digital requirements refer to the conversion of the captured image into a digital record. Finally, format requirements refer to additional or conditional required fields and in a Type-10 record or NIST transaction.

Note that the set of requirements applies to all poses of a subject.

*** Scene Requirements**

<u>Number of photographs</u>

<u>Level 50 and 51 records shall include at least five photographs of the subject: frontal, left and right profile, and left and right half profile. A half profile view consists of a face with a yaw pose angle of +- 45 degrees, and with Pitch and Roll angles of zero. Field 10.025, subject pose angle (SPA) values shall be recorded as (45, 0, 0) subject facing left, and (-45, 0, 0), subject facing right. Note that for half profile, the orientation of the head is rotated 45 degrees to half profile. (The rotation of the body shall not be required), and care should be taken to prevent the subject from keeping the head fixed while changing only the gaze. In addition, for half profile and profile photographs, the ear should be visible, pushing the hair back to the extent possible. For full profile images, the entire body shall be rotated with the head.</u>

<u>The "Head and Shoulders" photo composition</u>

The composition consists of a subject's head, partial shoulders, and plain background. For a frontal-facing pose, the width of the subject's head shall occupy approximately 50% of the width of the captured image. This width shall be the horizontal distance between the mid-points of two imaginary vertical lines. Each imaginary line shall be drawn between the upper and lower lobes of each ear and shall be positioned where the external ear connects to the head. A template and an example is shown in **Figure 1**. For other poses, the composition shall be rotated about an imaginary axis extending from the top of the head though the base of the neck.

This composition is applied to SAP levels 30, 40, and 50.
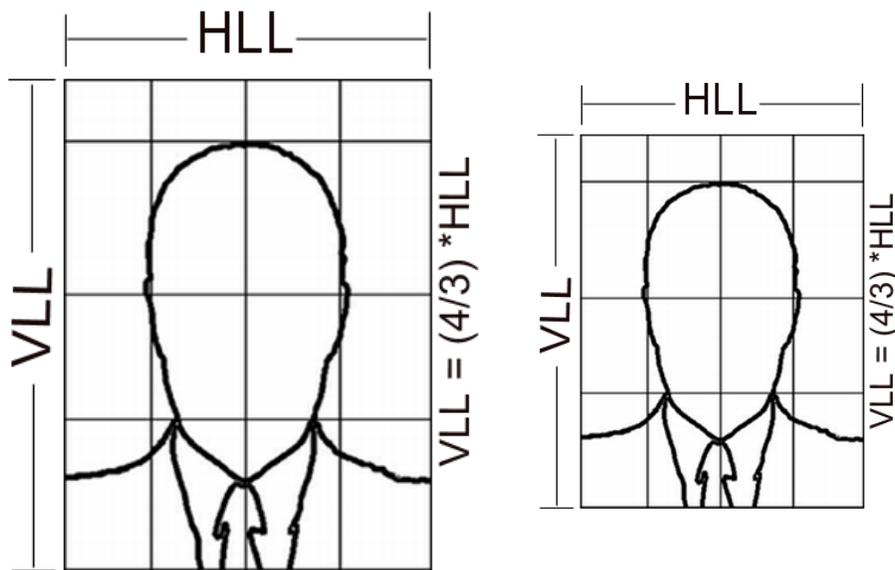
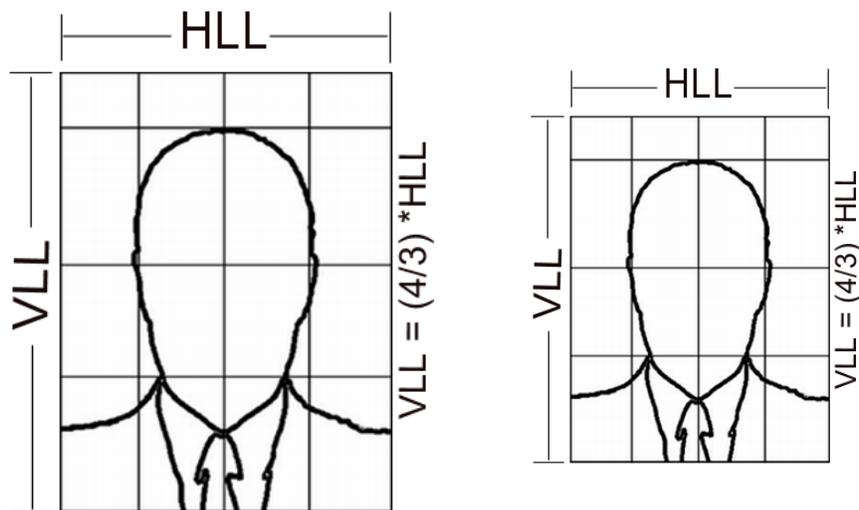| a. A template of the "head and shoulders" photo. The width of the head is ½ the width of the photo.	b. An example "head and shoulders" photo.

Figure 1– A facial image template and an example image that meets "Head and Shoulders" (levels 40 and 50) scene constraints.

<u>The "Head Only" photo composition</u>

The composition consists of a subject's head, and a plain background. For a frontal-facing pose, the width of the subject's head shall occupy approximately 70% of the width of the captured image. This width shall be the horizontal distance between the midpoints of two imaginary vertical lines. Each imaginary line shall be drawn between the upper and lower lobes of each ear and shall be positioned where the external ear connects to the head. A template and an example are shown in **Figure 2**. For other poses, the composition shall be rotated about an imaginary axis extending from the top of the head though the base of the neck.

This composition is applied to SAP level 51.

a. A template of the "head only" photo. The width of the head is 7/10th the width of the photo.	b. An example "head only" photo.

Figure 2– A facial image template and an example image that meets "Head Only" (level 51) scene constraints.

<u>Head centering</u>

For the frontal pose, the face shall be positioned to satisfy all of the following conditions:

The approximate horizontal midpoints of the mouth and of the bridge of the nose shall lie on an imaginary vertical straight line positioned at the horizontal center of the image.

An imaginary horizontal line through the center of the subject's eyes shall be located at approximately the 55% point of the vertical distance up from the bottom edge of the captured image.

For non-frontal pose, the subject shall satisfy these conditions when the head is rotated about an axis though the head and torso from the current pose back to center (zero angles) pose.

Visibility of Ears

The ear(s) shall be visible in frontal, profile and angled views for both "Head and Shoulders" and "Head Only" scene compositions. The hair shall be pushed back or tied behind the ears when appropriate.

From the 50/51 level description:

If hair covers the ears, then when possible, two photographs should be captured – one with hair in its normal state, and one with hair pulled back behind the ears.

Facial expression

The expression should be neutral (non-smiling) with both eyes open normally (i.e. not wide-open), and mouth closed. Every effort should be made to have supplied images conform with this specification. A smile with closed jaw is not recommended.

Eyeglasses

For subjects who normally wear eyeglasses, every effort should be made to capture the mugshots with the glasses on. If significant glare in the glasses is evident in the photograph, then a second frontal mugshot image should be captured of the subject without glasses. Specification of eyeglasses in the SXS field is required.

<u>Eye patches</u>

The wearing of eye patches is allowed only for medical reasons. In these cases, the specification of the patch, in the SXS field is required.

<u>Background</u>

The subject whose image is being captured shall be positioned in front of a background which is 18% gray with a plain smooth flat surface. A Kodak or other neutral gray card or densitometer shall be used to verify this 18% gray reflectance requirement.

The boundary between the head and the background should be clearly identifiable about the entire subject (very large volume hair excepted). There should be no shadows visible on the background behind the face image.

* Photographic Requirements

<u>Depth of field</u>

The subject's captured facial image shall always be in focus from the nose to the ears. Although this may result in the background behind the subject being out of focus, this is not a problem[1]. It is recommended that auto-focus on the central part of face be used with digital camera photography.

<u>Subject lighting</u>

Lighting shall be equally distributed on the face. There shall be no significant direction of the light from the point of view of the photographer.

The region of the face, from the crown to the base of the chin, and from ear-to-ear, shall be clearly visible and free of shadows.   In particular, there shall be no dark shadows in the eye-sockets due to the brow and the iris and pupil of the eyes shall be clearly visible.

---

1.   [1] For optimum quality of the captured mugshot, the f-stop of the lens should be set at two f-stops below the maximum aperture opening when possible.

Subject illumination can be accomplished using three point balanced illumination sources. A single bare "point" light source, such as a camera flash, is not acceptable for imaging.

Appropriate diffusion techniques shall be employed to eliminate hot spots on the facial image. These hot spots usually appear on reflective areas such as cheeks and foreheads.

Background lighting

Proper lighting shall contribute to the uniformity of illumination of the background, and the background shall be free of shadows.

Exposure calibration

The exposure shall be keyed to the background. Several areas of the recorded 18% gray background shall be used to verify the proper exposure. The averages of the 8-bit Red, Green, and Blue (RGB) components within each area shall be calculated. Each of the RGB means shall fall between 105 and 125 with a standard deviation of plus or minus 10. Furthermore, for every area examined, the maximum difference between the means of any two of the RGB components shall not exceed 10.

No saturation

For each patch of skin on the person's face, the gradations in textures shall be clearly visible. In this sense, there will be no saturation (over or under exposure) on the face.

No unnatural color or "red-eye"

Unnaturally colored lighting (e.g. yellow, red) is not allowed. Care shall be taken to correct the "white balance" of image capture devices. The lighting shall produce a face image with natural looking flesh tones when viewed in typical examination environments. "Red-eye" is not acceptable.

No color or grayscale enhancement

A process that overexposes or under-develops a color or grayscale image for purposes of beauty enhancement or artistic pleasure is not allowed. The full spectrum shall be represented on the face image where appropriate. Teeth and whites of eyes shall be

clearly light or white (when appropriate) and dark hair or features (when appropriate) shall be clearly dark.

<u>Distortion and angle of view</u>

No unnatural radial distortion of the camera lens, resulting in a diagonal angle of view of approximately 20 to 28 degrees

Fish eye effect, a type of distortion where central objects of the image erroneously appear closer than those at the edge, typically resulting in what appear to be unusually large noses in the image, is not allowed. While some distortion is almost always present during portrait photography, that distortion should not be noticeable by human examination.

For a typical photo capture system with a subject 1.5 to 2.5 meters from the camera, the focal length of the camera lens should be that of a medium telephoto lens.   For 35 mm photography this means that the focal length should be between 90 mm and 130 mm.   For other negative formats/sensors the recommended focal length is 2 to 3 times the diagonal of the negative/sensor

* Digital Requirements

<u>Pixel aspect ratio</u>

Digital cameras and scanners used to capture facial images shall use square pixels with a pixel aspect ratio of 1:1

<u>Image aspect ratio</u>

The Width:Height (i.e., HLL:VLL) aspect ratio of the captured image shall be 3:4. This corresponds to commonly used format sizes such as 600 x 800, 768x1024, etc., allowing for a COTS digital camera to be used for capture.

<u>No interlacing</u>

Interlaced video frames are not allowed and all interlacing must be absent (not simply removed, but absent).

Interlaced video frames are not allowed and all interlacing must be absent (not simply removed, but absent).

No digital zoom

The use of digital zoom (interpolation) to achieve specified resolution associated with Subject Application Profiles.

Image Compression

Non-frontal facial images shall be compressed using JPEG 2000 (JPEG is not allowed) meeting the maximum compression limits specified below.

There shall be one frontal facial image compressed using lossless JPEG 2000. If multiple frontal images are in the transaction, then one image must be compressed via lossless compression and the others can be compressed either using lossless JPEG 2000 or lossy JPEG 2000 that meets the maximum compression limits specified below. The best practice is to apply the lossless compression to the frontal image meeting the "Ears Visible" constraint.

The maximum compression ratio for both JPEG and JPEG 2000 of a rectangular region containing any exposed skin of the face, from crown to chin and ear to ear, shall be at most 15:1. This requirement is derived from studies of face algorithm matching at high and low resolutions. The non-facial portion of the mugshot, as well as the SMT Type-10 record, can be compressed up to a ratio of 120:1.

For JPEG, reference [14] provides source code to implement compression with both ROI and fixed compression ratios. For JPEG 2000, these capabilities are built into the implementation.

For both JPEG and JPEG 2000, care must be taken to account for automatic compression by camera hardware. Multiple compression stages can damage the quality of photographic data. When possible, minimum compression (highest

resolutions) should be applied at the camera level when external software performs the final (15:1 or less) compression stage.

The table on the following page provides the typical size of a single facial photograph using the compression recommendations contained in this section for levels 30 and 40, 50, and 51. We assume that the image is formatted as RGB888 (8 bits per color channel per pixel). For levels 30 and 40, we also estimate that since the face width is 50% of the image width, then the area taken by the face is 25% of the total image area. SAP levels 50 and 51 include the constraint of lossless compression for the frontal pose facial image as discussed above.

Allowed color space

A full color image shall be captured. To ensure that color images exchanged between differing systems can be correctly displayed or printed, images shall be converted to the device-independent color space, *sRGB*.

Table 1: **Example File Sizes after compression**

Level	Minimum WxH	Uncompressed Size (RGB888)	Size @ 2:1 Lossless Compression	Size @ 15:1 compression for the entire image	Size @ 15:1 compression for the face and 120:1 for the background
30	480x600	844 KB		58 KB	19.34 KB
40	768x1024	2.3 MB		156 KB	52.8 KB
50	3300x4400	42.5MB	14.2 MB		
51	2400x3200	22.5 MB	7.5 MB		

* Format Requirements

Subject Pose (POS) and Subject pose angles (SPA)

One of either the POS or SPA fields shall be used to denote pose angles.

The POS field code values "F", "R", and "L" can be used for images in which the Pitch and Roll angles are 0 and the Yaw angle is 0, 90, and -90 respectively. (The sign of the Yaw angle in the previous sentence corresponds to the field 10.020 where a right profile is when the subject turns to the left).

The SPA field 10.025 can be used for the above poses and shall be used for all other angled poses. The POS field 10.020 shall then be of type code "D", for determined 3D pose, instructing the user to use 10.025 as the reference for pose angles. (For example, a ¾ profile capture would require a POS field entry of "D" with the angle specified in the SPA field.)

In all cases, the uncertainty in the Yaw pose angle determination shall be less than 5 degrees of the frontal photograph, and 10 degrees in the non-frontal photographs. The uncertainty in the Pitch and Roll angles shall be less than 5 degrees.

Subject facial description (SXS)
The Subject facial description field shall be present in the transaction when one or more of the facial attributes given by the type codes of 10.026 is present in the image.

Subject hair color (SHC)
The Subject hair color field shall be present in the transaction. The code "UNSPECIFIED" for this field is not allowed.

Subject eye color (SEC)
The Subject eye color field shall be present in the transaction. The code "UNSPECIFIED" for this field is not allowed.

Face-Pose Values
(Informative)

* The definition and range of pose angles:

The Yaw and Roll angles shall be measured from the full face pose position and have a range of values from -180 degrees to +180 degrees. The Pitch angle shall have a range of values from -90 degrees to +90 degrees. The pose angle set is given by Tait-Bryan angles.

Yaw angle: rotation about the vertical (y) axis. A positive Yaw angle is used to express the angular offset as the subject rotates from a full-face pose to their left (approaching a right profile). A negative Yaw angle is used to express the angular offset as the subject rotates from a full-face pose to their right (approaching a left profile).

Pitch angle: rotation about the horizontal side-to-side (x) horizontal axis.

Roll angle: rotation about the horizontal back to front (z) axis.

The angles are defined relative to the frontal view of the subject, which has angles (0, 0, 0) as shown in **Table 1**. Examples are shown in **Figure 4**.
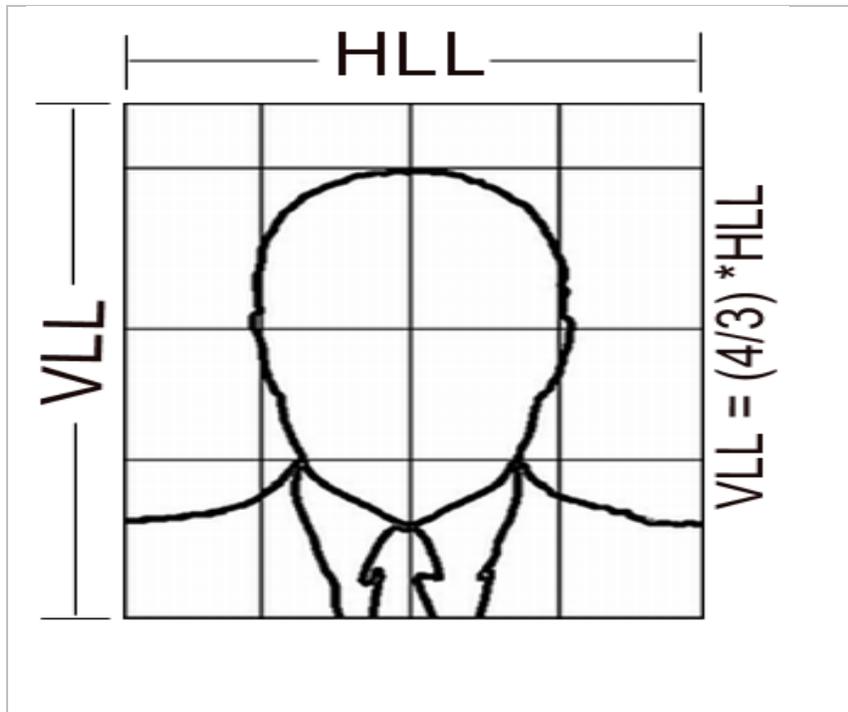
Figure 3 – The definition of pose angle set is with respect to the frontal view of the subject.
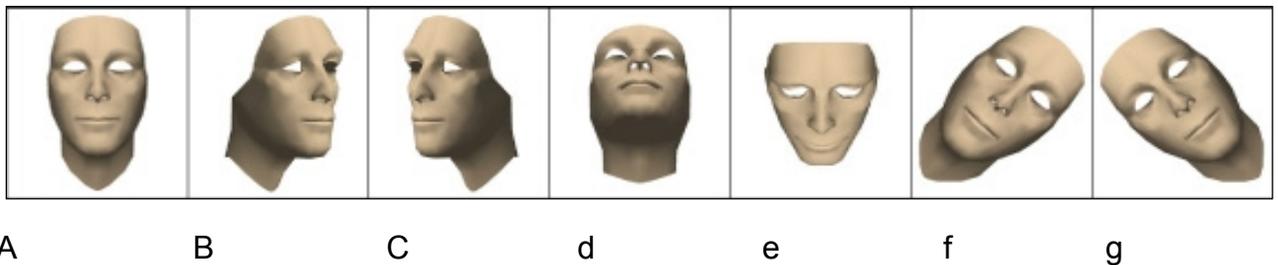

| A | B | C | d | e | f | g |

Figure 4– Examples of pose angles and their encodings. The pose angles (Y, P, R) of Figures (a) – (g) are given by (0, 0, 0), (+45, 0, 0), (-45, 0, 0), (0, -45, 0), (0, +45, 0), (0, 0, -45), and (0, 0, +45), respectively.

The uncertainty in the pose angles is given by the range 0 to 90, inclusive. It shall denote approximately a maximum value of possible deviation in the measurement of the pose. This shall correspond to a two standard deviation confidence interval.

The encoding of angles is in ASCII format, with the minus sign "-" used to denote a negative value and the plus "+" sign optionally used to denote a positive value. Pose angle uncertainty angles always are positive.

* The order of rotation through pose angles

As order of the successive rotation around the different axes does matter, the encoded rotation angle shall correspond to an order of execution starting from the frontal view. This order shall be given by Roll (about the front axis), then Pitch (about the horizontal axis) and finally Yaw (about the vertical axis). The (first executed) Roll transformation will therefore always be in the image (x, y) plane. Examples are shown in **Figure 5**.

From the point of view of executing a transformation from the observed view to a frontal view, the transformation order will therefore be Yaw, Pitch, and then Roll. Note however that the encoded angle is from the frontal view to the observed view.

(Y, P, R)=(0, 0, +45)	(Y, P, R)=(0,-30, +45)	(Y, P, R)=(-45, -30, +45)
A	B	C

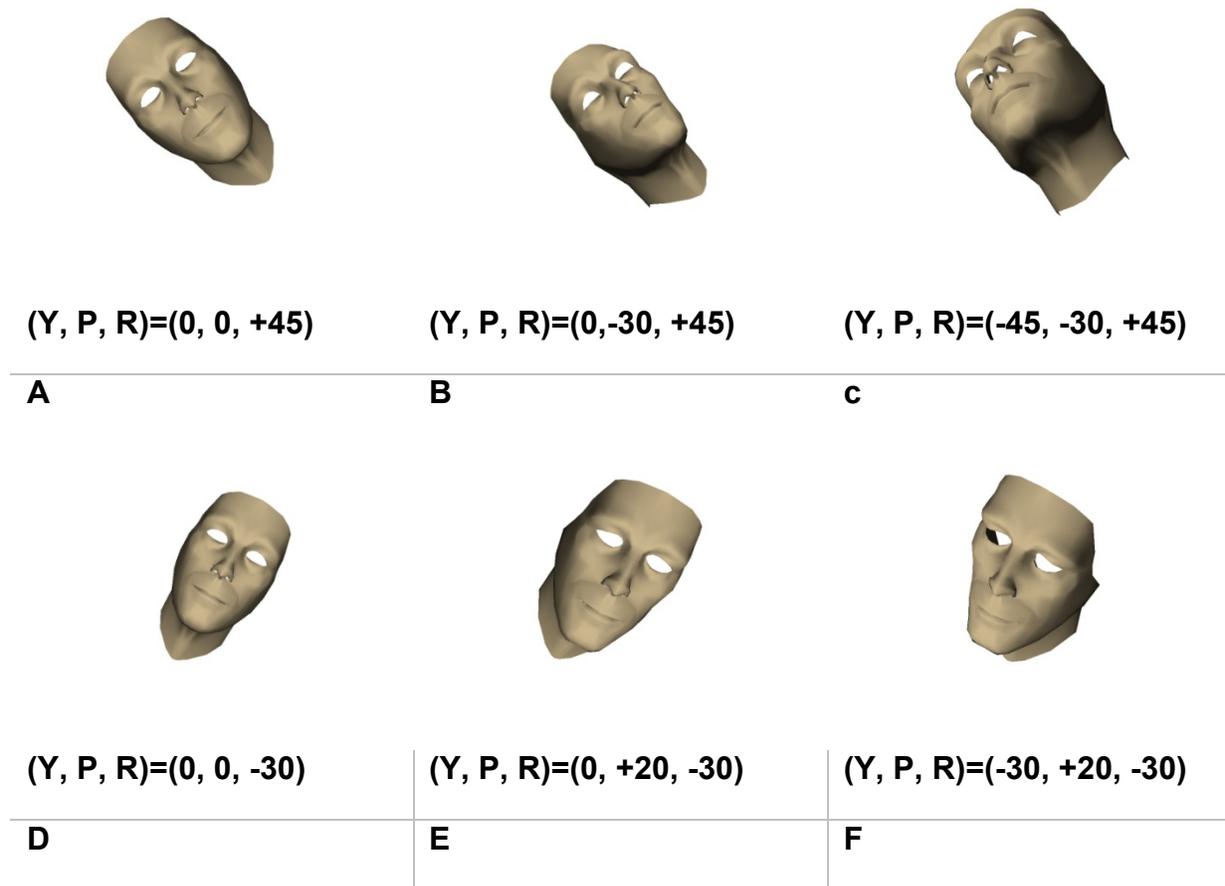| (Y, P, R)=(0, 0, -30)	(Y, P, R)=(0, +20, -30)	(Y, P, R)=(-30, +20, -30)
D	E	F

Figure 5 – Examples of the order of rotation through pose angles with an origin of coordinate system at the nose tip. Figures (a)-(c) show three successive rotation steps to achieve the pose angles (Y, P, R) of (-45, -30, +45). Figures (d)-(f) show three successive rotation steps to achieve the pose angles (Y, P, R) of (-30, +20, - 30).