False Rejects in Machine-Vision-Based Automatic Inspection Systems

The topic of false rejects is examined first with respect to achievement of performance objectives, and then with respect to specification and project administration. The relation of technical, human and specification factors to this topic is examined. FSI has been a manufacturer of factory automation products and systems since 1959 and provides related engineering, scientific and management consulting services. This paper was produced in response to numerous requests for guidance on this topic.

By Fred Turek and W.S. Tobey, FSI Machine Vision, Lombard, IL

Pass/fail type inspection is one of the most common applications for machine vision. The realistic goal of most automatic inspections systems is to pass zero bad parts while passing the highest feasible percentage of good parts. This means that the system is never allowed to pass a bad part, and allowed to at least occasionally fail a good part. The latter condition is generally referred to as a "false reject". While the working definitions of the term "false reject" vary widely, the general theme of all of them is the undesirable rejection of a good part by an inspection process.

This paper is divided into roughly 2 sections:

Understanding and Reduction of False Rejects This section uses a vague, general definition of "false rejects" for more effective coverage of these mutual goals.

Definitions, Specifications, and Expectations Establishes a framework for dealing with these aspects of the topic, and then does so.

Terminology Notes

Most definitions of the terminology used in this paper either self-evident, or are themselves the subject of the paper. Here are the working definitions used on a few others:

Customer The receiver of the machine vision solution. This may be in the traditional context where the customer is a different company than the provider, or in an "internal customer" context where this means the receiver of a solution provided by another person or entity of the same company.

Provider The provider of the machine vision solution, whether they are a different company, or a person within the same company as the customer. This may include persons or entities new to machine vision operating under a "technical umbrella" (with all tough problems solved) such as FSI Machine Vision's APS™ (Assured Path to Success™) program, an individual expert, or a firm that is providing a turnkey solution. Turnkey solutions may be some type of a complete manufactured "inspection machine", or a system added to already-existing production or material handling machinery. Approaches where there is no provider of the solution (e.g. where it is presumed that one of the pieces of equipment "does" the solution etc.) are not covered or considered to be viable.

"Fog of Vision" This is the Machine Vision version of Clausewitz's "fog of war" concept. We use this analogy to refer to the fact that, in most vision situations, a combination of technical, logistic and human factors and limitations makes it hard or slow to see, know and determine exactly what needs to be known. Examples of impacts are on seeing whether or not a vision solution is technically sound, whether a problem exists, clarifying perceived problems, identifying it source and testing potential solutions. This fog is a fundamental parameter of vision; and minimizing it is one on the essential tasks of sure fire success in a reasonable time frame.

Machine Vision Solution A combination of human actions, expertise and vision equipment that accomplishes a machine vision application.

Costs of False Rejects

The cost of a false reject is relevant to related decision making. This cost can vary from near-zero to very expensive. Some of the underlying variables are:

- **Does a reject shut down or disrupt the production process?**
- **What happens to a reject?** Is it scrapped, sent for more expensive retesting, or recycled for economical retesting?
- **What is the net cost to scrap this item at this stage of production?** (cost minus net salvage value)
- **Does a reject trigger additional consequences?** Examples are regulatory issues, special QA procedures, recordkeeping etc.

Sources and Reduction of False Rejects

For this section we temporarily use a usefully vague definition of a false reject as simply the rejection of good part. This does not define "good" leaving it to be in the eye of the beholder.

The goal of a pass-fail automatic inspection system is to:

1. Pass good parts and
2. Reject bad ones.

So, one could say that all machine vision & problems show themselves as merely two symptoms:

1. Reject good parts (fall rejects) or
2. Pass bad parts

Conversely, the source of false rejects is the entire universe of machine vision project, people and technical issues.

Next, this paper embarks on funneling processes which successively identifies, addresses and then narrows out these causes:

1. The wide end of the funnel includes administrative and communication causes

2. The mid-section includes solution and technical implementation weaknesses and malfunctions

3. The narrower end deals with systems that are running well at their technical limits.

*At the end of (or within) each is a guess at the percentage of actual or perceived false rejects (on non-FSI projects) where the cause is best characterized by that section.

At the broadest end of the funnel are statements of false rejects that are passed along without verification. Conflicting agendas can come into play as can normal grapevine dynamics. An example of the former is origination with persons adverse to use of machine vision due to its labor saving potential. More commonly, the "grapevine" can cause information (or its meta data such as completeness or trustworthiness) to be inaccurately relayed, or prevent all-important two-way, verification type questions / communications. One tool useful for this issue is the ability of the vision unit to conditionally store large amounts of (failed part) images. Either way, any serious quantification of false rejects should include only verified false rejects. A practice that defines uninvestigatable or uninvestigated reports of false rejects as extant false rejects may itself be creating a different type of false rejects. **5%***

The next stage of the funnel passes through situations where the solution is clearly performing as its designer / implementer intended, but the customer feels that there are too many false rejects. If the difference of opinion persists after clarification of what's happening, it usually reflects a lack of a good, agreed upon specification covering the core performance parameters such as those covered in FSI publication # 2980-0033-01E. (this is our "VADS"....Vision Application Data Sheet). **5%***

The next case is where the solution, even if on-site implementation is done well, mistakenly identifies clearly good parts as bad. This may be a calculated economical decision, for example, to implement / test it only with common parts instead of all parts or, a mutual decision that "poor" performance is still worth doing. But if this is at a level where the solution isn't capable of doing a reasonably good job on the application, there is a more severe problem. This is usually due to insufficient expertise by the implementer combined with insufficient work done prior to proceeding. While these two things seem like simple vendor errors, in reality the sources are both vendors and customers, and they are swept into those errors by various defective norms of the (still immature) machine vision field. These issues range from commercial norms (such as trying to define all vendor work prior to system purchase as "sales" and not utilize any services at that stage) to common conversational / advertising practices that misstate and mislead the process. (such as equipment-centric impressions that the vision unit "does" the application). **35%***

The next is a solution which IS capable of doing a reasonably good job, but which is not running / set up properly. This may seem like a straightforward matter, but the "fog" is a major player in this process, and often it isn't. The "fog", if not minimized, could make a 30 minute fix take three months. Three months is seldom spent, in which case it ends up not really working very well. **35%***

Now we move past the items which cause 80% of false reject problems to the technical "core" of this topic. While only **20%*** of instances fall purely within this "core", this section includes substantial material that is applicable to the entire false reject topic.

This section narrows our discussion to false defects that are occurring on a "reasonably good" solution; one that is generally capable of accomplishing the application and where any major problems have already been avoided or solved.

The main foundation statement of this section is:

The realistic goal of most automatic inspections systems is to pass zero bad parts while passing the highest feasible percentage of good parts.

(This paper is useful for [but does not cover] the other rare case of deliberately allowing passage of some bad parts in order to reduce cost or false rejects)

System parameters can be always be adjusted to be sufficiently stringent to pass zero bad parts. The absurd extreme is a useful illustration of this available tradeoff…a completely blind or incapable vision solution could be set to still achieve “zero defects” by setting it to reject every part. While the reject bin would be overflowing, it would achieve the goal of passing zero defects. Machine vision solutions that are successively better than this boundary case are capable of passing successively higher percentages of borderline good parts while still being adjusted to pass zero bad parts. The “feedback loop” for making these tradeoff adjustments is always slow and foggy, but such decisions may be made and implemented. .

The numeric goal of the machine vision solution is to maximize the percentage of good parts passed when adjusted to reject all bad parts. This is a function of three factors:

1. Strength of the overall vision solution. This is the degree to which the vision solution accurately and reliably perceives or measures the situation, including its ability to do this under sub-optimal maintenance and conditions. Mathematically, this is the degree of correlation between the actual and “as perceived” conditions. This is the overall “capital” available for the tradeoffs described in the next point.

2. Stringency or looseness of the pass / fail type settings.

3. Actual statistical distribution of the measured attributes. A situation where most of the good parts are borderline will tend to create more false rejects than a situation where most of the good parts are not borderline.

Let’s presume that #2 will be done close to optimally, (set just stringently enough to assure passing no bad parts), and that #3 is not a controllable factor. Then #1 becomes the fundamental determinant of how well false rejects may be minimized.

In short, the primary way to reduce false rejects is to obtain/provide a better machine vision inspection solution. While this fact is important to understand, it isn’t yet very specific because it encompasses an entire universe of equipment, expertise, actions and topics, with a time span of initial inception through long term ownership.

At the general level, the most common cause of poor solution performance is “missing pieces” in the solution due to failure to take a holistic approach to the machine vision solution. For example, this may be thinking that is overly equipment-centric or programming centric. Another example is failure to rigorously define and learn the mission, conditions and variables of the application, or to not address support and supportability issues. A detailing of this is beyond the scope of this publication. An example of a roadmap to such a holistic approach is FSI Machine Vision’s Assured Path to Success™ program. More information on this is available in publication #’s 2980-0074-01 (the actual roadmap) and SDOC-057-01.(an overview) Electronic or paper forms to assist the process of comprehensively learning and defining the mission, conditions and variables are also available, an example being the FSI “VADS” (publication # 2980-0033-01E.)

A design goal that favorably influences this is development of what we term a “robust” solution. This is a strong solution that easily differentiates good from bad. This makes it less vulnerable to real-life variables and less-than-optimal conditions or settings. From a customer standpoint, the strategic level strategy is to choose strong equipment, expert people, and a holistic approach.

At the tactical “optimization” level the general strategy is to find the cause of the problem or “weakest link” in the capability of the solution and then fix or strengthen it. This means decisively learning (through the fog) the cause of any false rejects. This requires a system which has easily-used capabilities to support that investigation. For example, the ability to “lift the hood” and to see exactly what is going on in the inspection process. As a minimum this should include the ability to do “step-by-step” execution from a saved, triggered image from the actual process, with a full display of each step’s results. The ability to conditionally store many images of (only) failed parts is always valuable for this process, and, for many situations (such as occasional false rejects on a high speed process) a near-necessity.

For our current purposes, a useful classification of inspection tasks is:

A. Verify presence (vs. absence) of something which should be there.

B. Detect presence of something should not be there, and make a pass/fail decision based on the numerical magnitude of its size or severity.

C. Check the magnitude of a dimension or of some other numerical parameter of the product.

Let’s rearrange and restate these as:

1. Verify presence (vs. total absence) of something that should be there

2. Detect presence of something that should not be there.

3. Measure and evaluate the numerical magnitude of something that should or shouldn’t be there.

Next we’ll examine these individually in the “reasonably good solution” context of this section.

#1 Verify presence (vs. absence) of something that should be there

This is the least frequent source of false rejects on "reasonably good" solutions. . A robust solution is usually easy to create for this mission because it searches for an expected object in an expected location. (The "vs. total absence" wording means that any finer evaluations of present items are excluded from this group.)

#2 Detect presence of something that should not be there

This a common source of false rejects on "reasonably good" solutions. A typical example of this is analyzing a surface for defects. Probably the most common cause of defects within this area is the encountering of "OK" features on the surface that were unanticipated at the time that the solution was created, and which appear (to the system) to be defects.

An example of this is a check for pits (small voids) in what should be a smooth metallic surface. A general solution strategy might be based on lighting design that makes interruptions of the surface appear dark, and an uninterrupted surface to appear light. If this solution is compatible with the range of products, product conditions and site conditions, then it is typically one of the best solutions. Now, in our example, a situation occurs with loose chips of metal on the surface at the time of the inspection. While unanticipated, this is considered an acceptable condition for the products, yet they appear dark to the solution, causing the parts to be rejected.

The remedy for this is to start by methodically learning, communicating and planning for the range of products (and their variable conditions) to be inspected. (this is also covered by the FSI VADS) Since the cost of a good vision solution is roughly proportional to the range of scenarios, the other "CYA" extreme of saying, "be ready for every possibility" should also be avoided, as this would cause the most credible providers to escalate their prices or run for the hills, leaving only the others. Usually the best standard to use is "what has already been occurring?" and "what is likely to occur?"

The second most common cause of false defects in this scenario is a solution that is marginal in its ability to differentiate between defects and normal features. The answer here is the general one which is a stronger, more robust solution.

#3 Measure and evaluate the (usually numerical) magnitude of something that should or shouldn't be there.

This section is useful beyond it's apparent narrow topic because:

- Most presence and absence detections also utilize numerical criteria
- It provides logical and mathematical underpinnings useful for all false reject and numerical accuracy topics. .
- It mathematically illustrates the four-way equation between strength of the solution, reliability at rejecting all bad parts, and minimization of false rejects, and statistical distribution of good parts with respect to the relevant parameters.

Let's start with a very simple gauging example where a dimension of a part simply must be 1.000" +/- .1". (A crude tolerance was deliberately chosen for illustration purposes) By rigorous inspection logic, the "+/-" is actually two inspections, despite the ability to do them with a single software tool. Let's cut this down to one by ignoring the "minus" inspection, and restate the reduced mission as assuring that the part never exceeds 1.100". Now let's discuss the (required) accuracy of the measurement process by working towards a precise "pass/ fail" type statement of the mission. (Some of these topics are covered in more depth in our "gauging accuracy" paper.)

First, let's clarify the definition of accuracy as: *the maximum error that will occur in the overall measurement process.* This discards as meaningless the types of machine vision "accuracy" answers that are usually given to this question such as figures derived from the resolution of the imaging array plus (sometimes) conversion factors to spatial resolution and to reflect enhancement by sub-pixeling capabilities. This answer merely describes how much error is added (to the total error of the measurement process) by the granularity of the array. This is usually equivalent to presenting the strength of the strongest link of a chain as being the answer to a question about the chain's strength. Depending on the level of expertise of the respondent, this answer is either myopic or deliberately misleading.

The actual accuracy is based on the sum of all sources of error, the largest of which are usually unrelated to the imaging array of the machine vision unit. The desired level of accuracy in the measurement process has a substantial effect on cost, so it must be chosen carefully. For our discussion, let's postulate the following typical accuracies and costs for a turnkey solution to our mission:

+/- .1": $5k,
+/- .01": $13k
+/- .001" $30k,
+/- .0001" $70k,
+/- .00001" $300k.

Now let's try to state the pass/fail mission explicitly. Some of these deliberately repeat common errors or overkill choices.

- The first attempt might be to say: "reject every part that is over 1.1 inches, and pass every part that is less than 1.1 inches". This sounds fine but isn't.....it would inadvertently require a solution of infinite accuracy and infinite cost.....for example, one that could differentiate between a 1.100000001" bad part and a 1.0999999999" good part. Even the $300k solution would not be sufficiently accurate to accomplish this.

- Another would be to use the part tolerance as the answer to the measurement accuracy question. For example, "the required accuracy is +/- .1 inches " (The coarse part tolerance was chosen to make the ramifications of this clearer than is usual). In this case, measurement process accuracy of +/- .1" would *not* mean that the part can be up to .1" oversized, but that the "ruler" used to measure it may itself have/induce an error of up to .1". This method usually does not specify sufficient accuracy if there are borderline good parts.

- Common practical choices range from 2 to 6 times as accurate as the measurement process. Let's posit that accuracy of +/- .02 is required, and the +/- .01" is system is a good choice. One could play it "safe" and say that: "Vendor will not be fully paid unless accuracy of +/- .01" is proven" In this case, reputable vendors would have to include the .001"/ $30k system to be 100% sure of hitting .01", plus most would add a few $k to cover the proving process.

- State "Required accuracy of the entire measurement process to be +/- .02". This approach (state just what's needed) is usually a good approach.

For this example let's choose a measurement solution with an accuracy of +/- .01", which is ten times "better" than the part tolerance. Again, the mission is to assure that every part larger than 1.1" is rejected. The worst-case scenario regarding the measurement inaccuracy tending towards allowing a bad part to pass would be a (perceived) measurement .01" lower than actual. To assure that the largest possible measurement system error will not allow a bad part to pass, the system pass/fail criteria setting must be offset to allow for this error. And so it is set to reject all parts with a (perceived) measurement over 1.09" (1.10" - .01" = 1.09").

With respect to creation of false rejects, the worst case scenario is for the vision solution is to read .01" too high and perceive a part that is barely over 1.08" to be barely over the reject setting of 1.09". And so it is *possible* for any part between 1.08" and 1.10" to become falsely rejected. The width of this "possible false reject band" is twice the ± .01" (in)accuracy of the measurement process.

How many false rejects will this create? The primary remaining variable that determines this is the frequency with which "border line" parts occur. The 2 most extreme hypothetical scenarios illustrate this concept:

- If, a day's production consisted of 1.0950" parts, then the false reject rate would be 100%.

- If the day's production consisted of only 1.0000" very good parts and 1.15" very bad parts, then the false reject rate would be 0.000%.

Here are 2 more realistic examples of this:

- The inspection might be to watch for a tooling failure, where the parts are always "dead on" until the tool breaks, at which point they are "a mile" off. In that case, parts within the 1.08" – 1.10" "possible false reject" band would never occur, and there would be zero false rejects. A more accurate system would be a waste of money.

- An inspection is monitoring the accuracy of a machining process, where there may be many "borderline parts", and thus many false rejects. This case would require a more accurate vision solution, to reduce false rejects to an acceptable level.

This same concept may apply for applying numerical criteria for the defects rather than the product. For example, to reject a part if there are discolorations or pits greater than .01" in size. It is common for the intent of such values to only to be to define the "must reject" condition; i.e. to not be concerned about rejections where the defect occurs but is smaller than the defined criteria. However this should still be addressed via defining a looser "required accuracy of the measurement process" to avoid creating concern by or overkill quotations from credible / conscientious potential providers.

This concludes our "funnel" section and also the first "half" of this paper.

<u>Definitions, Specification and Expectations</u>

In the previous half of this paper we chose a vague definition of false rejects. Now we address the topic of a more precise definition.

A "Gold Standard"

As a foundation to tackling this topic, let's define a "gold standard" baseline scenario. "Gold" does not necessarily mean a good decision / plan, and in fact, following a few sections of this may be an overkill (=poor) decision. But this does provide a useful framework and starting point for good planning and practical decisions.

1. The application conditions, mission, and range of scenarios to be accommodated has been explicitly defined. This may be the *<u>end</u> point of substantial and significant work* which might include outside consultants and feasibility study work. (An example framework is a thoroughly completed version of our "VADS", publication # 2980-0033-01E) At this point the mission statement includes "must fail" provisions, but not "should pass" provisions. Mutual agreement has been obtained from all relevant parties that this document is sufficiently complete, precise & specific.

2. Add "should pass" criteria in the work under #1. These are typically the same format as the "must fail" criteria, but with *different* numerical values. Note that this is a modified value for the attribute, not a percentage of allowable false rejects.

3. The application as defined is determined to be feasibly doable by persons sufficiently expert to make that determination. In short, that the strength of the solution is greater than the difficulty of the application. This may be anything from (for a very simple application) a few minute process by an employee / provider of the customer to a large-scale feasibility study, run by a consultant or turnkey solution provider. It is important to note that the "difficulty of the application" is generally determined by the range of scenarios that must be handled, part presentation specifics, and constraints (or lack of constraints) on the solution rather than by the general nature of the inspection.

4. A sufficiently strong solution by a sufficiently expert provider is chosen.

5. A thorough and holistic approach to the application is followed and completed. (One example of such a roadmap is the FSI APS™ program.)

6. Some method of verification and review of apparent false rejects is put in place. This is usually easy to do when the system has strong tools for this purpose (such as the ability to *conditionally* store a large number of images of rejected products). Otherwise other methods to do this are put in place. Only verified instances of false rejects are classified as such.

7. If the mission includes numerically defined pass/fail criteria, the "required accuracy of the (entire) measurement process is specified and the principles from earlier in the paper are implemented to define a band of borderline good parts which, if rejected, are not defined as false rejects. Any evaluation is done via a more accurate measurement process, and it's (in)accuracy is also allowed for in the interpretation of the evaluation results. (#7 is actually an expansion of a topic already included in #1, #2 and #6)

In general, #1, #3, #4 & #5 are also essential to the general success of the project. A project can be a success without #2, #6 and #7; these are chosen for more thorough handling of false reject issues.

Practical Advice for the Customer

This paper's primary recommendation for minimization of false rejects is to emphasize a strong, relevant solution as a higher priority than efforts focused directly on false rejects. More specifically:

1. Thoroughly assess and define the likely conditions, range of products to be inspected, constraints and mission of the application. This should be broad enough to define the likely range of scenarios, but should avoid overly broad and vague descriptions which are counterproductive to obtaining the best providers by differentially causing "no-bids" or de-facto "no bids" (via. large escalations) by the most expert vendors, leaving only the others. Publications such as FSI's "VADS" (publication # 2980-0033-01E) provide a framework for doing this.

2. Choose a strong solution, this being a combination of people, equipment, companies and methods. Avoid "unit-centric" inquiries; structure them to encourage a holistic approach. Determine where some "overkill" regarding capabilities (of people, methods or equipment) is economical, and choose to include it in those cases.

3. Choose and implement one of the three approaches in the next section regarding false rejects.

Practical Specification Advice

This provides practical advice for the definition of false rejects and the specification of false reject performance.

Here are three approaches:

1. **No numerical treatment of the "false reject" topic and no explicit "should pass" specifications.** Emphasis is placed on #1, #3, #4 and #5 in the previous section, for projects where these alone will result in acceptable-to-all false reject performance if a reasonably good job is done on creating and implementing the solution.

   "Required accuracy of the measurement process" *is* specified for inspections with numerical criteria; in those cases, this is an important number that is relevant to both false rejects and general system capability. This can be done using the referenced "VADS" document # 2980-0033-01E.

   This does not provide a complete explicit framework for resolving "fine point" differences of opinion regarding false rejects. However, on the projects that this section is intended for, most false reject issues can and should be resolved by mutual verification a viable solution is in place and is functioning, or rework of the solution until this becomes the case.

2. **Fully implement the previously described "gold standard"** Set the "should pass" criteria so that it'

is realistic to achieve zero false rejects, but then allow for a very small percentage of false rejects. .

3. **Decide on a maximum false reject percentage for a fixed and sufficiently large sample set.** This requires use of the same sample set starting with just prior to final agreement on the specification. This removes the "product production" variable. In this case, the "should pass" criteria may be left to be the same as the "must fail" criteria.

Conclusion

Minimizing false rejects is one of the two fundamental goals of an automatic inspection solution, so sources of false rejects cover nearly the entire gamut of machine vision topics.

It *is* possible to specify and quantify the performance of a machine vision solution relative to false rejects, but specifying a false reject *rate* is usually not a technically sound way to do that. A framework for doing this is given. At the core of this is either specifying the accuracy of measurement processes (which underlie most inspections) or specifying different individual "should pass" values for those attributes.

False rejections are best minimized by developing a thorough but practical definition of the objectives and conditions of the application, and obtaining or implementing a strong machine vision solution. A strong solution includes strong holistic methods, expertise, architecture and equipment. Pervasive equipment-centric norms in the still-immature machine vision field tend to obscure the inherent importance of these items. The foundation and framework provided in this paper supports success with false rejects specifically and with machine vision applications in general.