VLBA Acquisition Memo # 107

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
HAYSTACK OBSERVATORY
WESTFORD, MASSACHUSETTS 01886

Area Code 617
692-4765

LEARNING FROM ERRORS

Hans F.Hinteregger
Haystack Observatory
January 21, 1988

1.0 NRZM

VLBA and MKIII longitudinal recording formats differ only slightly. On tape along a track both use the NRZM (non-return-to-zero mark) binary channel code representation in which a 'one' is encoded as a 'mark', in this case a magnetization reversal, and detected as a transition between binary (+/- unit) levels.

An inductive read head can sense only flux changes. This is why in magnetic recording the flux change is considered the 'mark' or physically active symbol.

The 'zero' in NRZM is a 'space', the physically passive no-change symbol. It is detected when no transition is sensed in a synchronized bit interval.

The qualitative physical yin/yang complementarity of ones and zeros in NRZM was kept in mind and is used to some advantage in the design of the VLBA/MKIII formats.

2.0 BIT ERRORS PAIRED

Bit errors in the NRZM representation are fundamentally paired. This is because flux changes (transitions, zero-crossings) strictly alternate in direction. It follows, for example, that a single transition (mark, one) somewhere between its nearest neighbors cannot be destroyed.

Misses and hits -- ones detected as zeros and zeros detected as ones respectively -- can appear only as closest partners of the same or opposite kind.

Closest, in this context, does not mean that bit error partners must necessarily occupy adjacent bit cells.

Rather, bit error partners of the same kind may be separated only by an unbroken string of spaces (zeros).

Anti-similarly, it is easy to show that bit error partners of opposite kind (hit/miss or miss/hit combinations) may be separated only by an unbroken string of marks (ones).

3.0 PARITY

VLBA/MKIII formats use simple longitudinal (in-track) byte parity. Bytes are nine bits long, parity bit last.

Odd parity is used everywhere except in the frame sync. Since there must be at least one mark in every byte a run limit of 16 zeros is thereby enforced.

4.0 SYNC WORD

The frame sync word consists of four consecutive 111111110 even parity bytes. Since any consecutive nine-bit segment of the sync word has even parity the sync word cannot appear without at least two parity errors in any bit-shifted position in any other part of a frame.

This property together with the expectation that tape flaws produce many more misses than hits -- destroy transitions more easily than create them -- permits a very simple format sync algorithm to be used safely:

Detection of the sync word unconditionally syncs or resyncs the format. Resync means that a sync word was found at an unexpected place; resync unambiguously detects and terminates the bit-slipped condition.

5.0 BIT SLIP

Bit-slip of one bit in either direction may occur 'legitimately' only as a result of a very rare, very long, very deep dropout.

Apparent slips of more than two bits are almost certainly a sign of a hardware problem such as oscillating amplifier or comparator or defective bit sync.

The presence of the bit-slipped condition, which persists -- is locked in -- after the flaw that caused it is gone until resync occurs, results in a 50% parity error rate during that interval.

A rather high frame validity threshold of order 255 parity errors in a VLBA or MKIII frame is adequate to meet the important systematic error spec that less than 10 per million bits labeled valid actually be out-of-sync, which might otherwise be violated if two independent opposite slips occur in the same frame.

6.0 ERROR DOUBLETS

In a binary world of transitions the statement "what goes up must come down" and the corollary "what doesn't go up can't go down", and vice versa, are absolute truth:

The corollary describes a pair of misses, mm. The mm is the most common atom-of-error because it represents loss of active response writing and/or reading transitions. In other words if a zero-crossing in one direction is missed the next zero-crossing in the opposite direction must also be missed.

The first statement describes a pair of hits, hh. If hysteresis in the transition detector (comparator) is larger than tape and interference noise then pairs of hits will be much less common than pairs of misses.

Tests by Bill Petrachenko on an early MKIIIA prototype indicated a miss-to-hit frequency ratio of about 5:1.

It's not clear that the ratio of signal to hysteresis level in MKIIIA or the VLBA prototype has been set optimally to minimize total error rate.

Nor whether efforts to reduce interference within the VLBA prototype have substantially increased the miss-to-hit frequency ratio as well as lowered the total error rate by reducing the hit rate more than the miss rate.

As indicated before, mixed error pairs are also possible, that is, miss-hit or hit-miss combinations, mh or hm. These represent transitions which have been grossly displaced so that they are detected in the next or previous bit-cell respectively if it is a space (zero).

The relative frequency of isolated mixed error pairs is thought to be very low since no obvious mechanism for producing them is known.

However, the bitslipped condition can well be described as a continuous run of hm or mh error pairs depending on the direction of slip.

A bit error rate tester capable of distinguishing and counting simultaneously mm, hh, and perhaps mh and hm bit error pairs as well as making a histogram of distances between them would be useful for investigating whether the channel can be improved electronically as well as to characterize tape flaws statistically. It may be possible to include such functions in the acquisition decoder.

7.0 PER = 2 x BER

The parity check will fail (indicate a parity error) only if an odd number of bits in the byte are in error. An isolated pair of bit errors will be detected only if the partners in crime are not both in the same byte. In that case they must be in adjacent bytes and two parity errors will result.

What fraction of bit error pairs result in parity error pairs? I'll show that the answer is 2/9 for isolated pairs of misses and should be about the same for the rarer error doublets.

This fraction, about 22%, is the theoretical bit error detection efficiency of the 9-bit odd parity NRZM code.

If we accept this for the moment and note that bit error rate is expressed in units of bit errors per bit while parity error rate is measured in units of parity errors per 9-bit byte, it is clear that parity error rate is statistically equal to twice the bit error rate.

8.0 DETECT 2 OUT OF 9

Assume an initial bit error, a miss, occurs with an equal 1/9 probability in any bit position of one byte. The probability that this error and its partner will be detected (result in two parity errors) is equal to the probability that the next mark which will also be missed is in the next byte.

Odd parity guarantees that at least one mark per byte is encoded. This means that the probability that the next mark is in the next byte is equal to the probability that it is not in the remainder of the present one.

For a random bit string the probability that any given bit has a particular value is 1/2. The probability that the n bits which follow the first miss and are part of the same byte are all zeros is 1/2 to the nth power.

If the first miss is in the last (parity, n=0) bit position the next (partner) mark must be in the next byte and will also be missed. The probability that each miss yields a parity error is unity.

This is true only under the assumption that there are no (odd numbers of) additional errors in the second byte, in other words that the error doublet is sufficiently isolated.

The assumption above is empirically justified even in the middle of most dropouts in which errors tend to occur in bursts that span several hundred bits. An overwhelming majority of parity errors can be observed (on a suitably triggered scope) to occur in adjacent pairs.

This implies that the mean distance between error doublets is probably at least an order of magnitude larger than the mean distance between partners. The latter is theoretically two bits. One can conclude that most dropouts appear gradual and shallow to the present bit detection circuit.

If dropouts were intrinsically sharp and deep (or engineered to appear more that way, by for example, increasing comparator hysteresis), so that they produced mostly unbroken runs of misses (erased all marks), then the odd parity check would be a nearly perfect logical dropout detector since an odd number of misses would then occur in every byte wholly within the 'boxcar' dropout.

For an initial miss in the first bit position in a byte (n=8) the probability that all eight remaining bits in the byte are zeros is not 1/256 but 1/128 since in this case alone the parity bit is determinate. The sum of these probabilities for all nine bit positions is therefore not 511/256 but 2. So the average probability that a pair of misses results in an adjacent pair of parity errors is 2/9.