ATOM: Super Compact and Flexible Format to Store and Transmit GNSS data

Igor Artushkin, Aleksey Boriskin, Dmitry Kozlov
Magellan, Russia

BIOGRAPHY

Igor Artushkin has been working in Magellan since 1994 as a Software Engineer. Since 2002 he is Senior Software Engineer. He received his MS in Mathematics and Mechanics from Moscow State University.

Aleksey Boriskin has been working in Magellan since 2005 as a Software Engineer. He received his MS EE degree from Moscow Aviation Institute. As a post graduated student he is additionally working for his PhD in GNSS data processing area.

Dmitry Kozlov has been working in Magellan since 1993 as a Senior Scientist. Since 2002 he is Algorithm Group Manager. He received his MS EE degree from Moscow Aviation Institute and his PhD in Signal Processing Theory from the Institute of Automatics, Moscow.

ABSTRACT

GNSS data (observables, corrections, positioning results etc) are usually stored in some memory for post processing and/or sent through some data link for various real time applications. The problem of compact presentation of GNSS data can be very important in some cases. For example, some low band data links do not allow sending RTK observables/corrections at 1 Hz using standardized protocols like CMR and RTCM-3. This problem is even more important when data from more than one GNSS are to be transmitted (e.g. GPS+GLONASS). Another example is GPRS/NTRIP when data link has usually sufficient bandwidth, but user pay for the actual traffic. The more compact protocol is used, the lower expenses user has. So developing compact data formats which maximally effectively utilize data link bandwidth is very important task.

The other issue related with data storage/transmission is growing GNSS. More and more GNSS appear and will appear in future. Each of them provides variety of signals and corresponding observables. So it is desirable to have some universal presentation format which can effectively serve any combination of GNSS and their particular observables.

Given paper describes new ATOM™ format from Magellan for data storage/transmission. ATOM is organized as universal format capable of generating the following group of data: GNSS corrections, GNSS raw observables, receiver positioning results and attributes, Satellite navigation information. The most important groups are GNSS observables and corrections. ATOM allows different presentation options for these data, from most compact to most full.

The paper gives the overview of ATOM format and provides throughput figures in comparison to existing formats.

INTRODUCTION

Multiple GNSS, multiple bands and multiple signals have entered our life and our receivers. Many legacy data formats (standardized and proprietary) cannot support the new reality well. That is why GNSS community spends a lot of energy to develop GNSS data presentation formats to make them 'ideal' to support current reality and future forecast.

On the other hand, a lot of novel applications require either long time data collection with high update rate (for office s/w applications), or real time data transmission via narrow band data links (for RTK operation). This dictates that GNSS data formats must be maximally compact not losing at the same time final performance.

RINEX-3 [1] and RTCM-3 [2] status and activity reflects the progress with new data formats matched the new reality. The hard work has been already done and even harder work is to be done in future.

At the same time, almost each particular vendor wants to have its own data format which fits its own specific applications better than existing standards. The reality one must accept is that proprietary formats will exist for very long time regardless the availability and maturity of standardized messages.

Magellan also has developed its own proprietary binary data format ATOM. The name stands for Adaptive Transmission of Optimized Messages and emphasizes the main distinguishing ATOM feature: possibility to present the data in compact form. ATOM is open for further extensions with new messages or update already existing messages (ATOM version number is provided for each message). We have foreseen a good deal of reserved bits to allow ATOM extension in future. All the ATOM fields are not obligatory aligned by integer bytes boundaries; however for extra convenience we group some fields to fit the integer number of bytes.

ATOM has been integrated into latest Magellan GPS+GLONASS L1&L2 products:

- ProMark™ 500 land survey receiver and corresponding office s/w
- MB 500 OEM board

The key features of ATOM include:

- Outputting the widest variety of GNSS data with any update rate
- Supporting different customization options from maximally compact to maximally full
- Being in line with existing RTCM-3 and NMEA-3 messages as well as RINEX-3
- Universal presentation form for different GNSS data
- Possibility to use ATOM for raw data recording and as differential protocol

Given paper gives the overview of ATOM features and algorithms behind them. At the same time we do not provide the detailed description of ATOM protocol.

The paper is organized as follows:

- Section TRANSPORT describes transport layer used to generate ATOM messages
- Section OVERVIEW gives high level look to six primary ATOM messages
- Section OBSERVABLES provides more detailed description of presentation of raw data and corrections in ATOM
- Section THROUPHPUT demonstrates the bit consumption of ATOM observation messages when using different presentation options
- In CONCLUSION we enumerate in compact form the distinguishing ATOM features which make ATOM ideal for various Magellan applications

TRANSPORT

While ATOM is proprietary message, it uses standardized RTCM-3 transport layer. This decision was made to allow in future any 3rd party vendor to decode ATOM using standardized RTCM-3 decoders.

RTCM-3 messages are varied in numbers from 1001 to 4095. Numbers 4001 to 4095 are reserved for proprietary usage. Each vendor can ask RTCM to assign a unique number from this range to be used exclusively by its own data. The number 4095 is reserved for Magellan [2].

As a result ATOM transport layer is the same as any standardized RTCM-3 message [2]:

Data field	Value/range	Notes
Preamble	11010011	8 bits, fixed
Reserved	000000	6 bits, for future use
Message length	0…1023	10 bits, length in bytes
Message	0...1023 bytes	Any RTCM-3 data of variable size
CRC	24 bits	QualComm definition CRC-24Q

Table1. ATOM transport layer

In turn, presentation of each ATOM message looks as follows:

Field	Value/range	Comment
Message number	111111111111	12 bits, fixed for Magellan ATOM
Message sub-number	0000…1111	4 bits, to distinguish between ATOM groups
Message version	000…111	3 bits, to allow firmware upgrades in future
Message body	0… 8165 bits	Data themselves

Table2. Generalized ATOM presentation form

Each particular ATOM message fits the presentation above.

Additionally to support easier parsing, we allow any ATOM message (including transport) to be encapsulated inside legacy Magellan (Ashtech) $PASHS frame as follows:

$PASHS,***,<atom_length><atom_data>cc<CR><LF>

where:

*** stands for any of COR/MES/PVT/ATR/NAV/DAT;
atom_length is 2 bytes atom_data length (in bytes);
atom_data is original ATOM data;

cc is 2bytes checksum;
<CR><LF> is carriage return line feed combination.

OVERVIEW

At the moment, ATOM supports six primary groups of GNSS data. ATOM is open to add more groups in future to the currently supported list. The short overview is given in the table below:

Group type	Group ID	Counterparts
GNSS corrections	4095,1 or ATOM, COR	RTCM-2 20,21
GNSS observables	4095,2 or ATOM,MES	RTCM-3 1001-1004, 1009-1012
Positioning results	4095,3 or ATOM,PVT	NMEA-3 GGA, GST,GSV etc
Receiver attributes	4095,4 or ATOM,ATR	RTCM-3 1029, 1033
Navigation information	4095,5 or ATOM,NAV	RTCM-3 1019, 1020
Raw GNSS data frames	4095,6 or ATOM,DAT	N/A

Table3. Primary ATOM groups

Groups MES and COR refer to GNSS observations. MES satisfies general RTCM-3 requirements [2], COR satisfies general RTCM-2 requirements [3]. They are presented in very similar form and can be converted to each other provided reference position and ephemeris data. These data can be converted to (generated from) standardized RTCM messages and/or RINEX files.

Group PVT allows outputting positioning results such as position, velocity, time, Satellite tracking/usage status. Additionally it contains the information about position latency and accuracy. These data can be converted to (generated from) standardized NMEA-3 messages.

Group ATR allows generating receiver/antenna attributes; say receiver name/serial number/firmware version and/or antenna name/serial number. It also allows specifying antenna reference point against survey point as well as any user defined message generation.

Group NAV allows generating navigation data extracted from GNSS data streams. NAV supports generating GPS, GLONASS, SBAS ephemeris and almanac as well as some other valuable information like broadcast GPS ionosphere parameters. GPS and GLONASS ephemeris messages are the copies of standardized RTCM-3 messages 1019 and 1020.

Group DAT allows generating raw navigation data stream (frames) decoded from any GNSS signal receiver tracks. Currently it supports only GPS, GLONASS, SBAS data extracted from L1CA signal, but DAT is open to generate the data from any other signal, e.g. GPS L2C.

Each group contains a number of particular messages/blocks which can be optionally enabled or disabled. Each group has its own default configuration.

Consider for example in more details the organization of ATOM,PVT message. It starts with a header (10 bytes) which contains the following data:

Field	Comment
Message number	111111111111 = 4095
Message sub-number	0011 = 3
Message version	001 = 1
Multiple message bit	1 indicates that more 4095,3 message(s) will follow for the same time tag
The number of blocks	Specifies how many data blocks follow
Position engine ID	Clarifies position engine configuration
The number of Satellites	The number of Sats: Potentially seen Tracked Used in position
Primary GNSS system	Defines the meaning of time tag and position datum
Time tag	
Reserved bits	For future use

Table4. ATOM,PVT header data

Currently the following PVT data blocks are supported.

Block type	Block ID	Size, bytes
Position	COO	25
Velocity	VEL	11
Clock	CLC	9
Accuracy	ERR	8
Latency	LCY	2
Attitude	HPR	10
Baseline	BLN	15
Miscellaneous	MIS	22
Range Residuals	PRR	2+4*Nsat
Satellites status	SVS	2+4*Nsat

Table5. Supported ATOM,PVT blocks

ATOM,PVT is open to add more blocks in future. It should be also noted, that currently all PVT data are output under the same header (possibly with unique update rate for each block), i.e. inside single ATOM,PVT transmission. At the same time, each particular block (e.g. COO or SVS) can be potentially output under its own header, i.e. using separate ATOM,PVT transmission. In latter case, multiple message bit in ATOM,PVT header is set accordingly to compile complete position epoch data from different transmissions.

OBSERVABLES

Groups MES and COR are the most variable and modular compared to other ATOM groups. They are usually the most interesting for end user, because they:

- Provide raw data output for 3rd party real time applications which can work with Magellan receivers
- Provide raw data recording for further post processing in the office (Magellan and 3rd party applications)
- Generate data using data link to perform RTK operation between base and rover receivers

A. GENERAL ORGANIZATION

The key features of ATOM observation messages are as follows:

- Supporting GPS, GLONASS, SBAS ranging data, open for Galileo and other GNSS
- Using RINEX-like signal and observables naming convention
- Generate up to pseudo-range (C), carrier phase (L), Doppler (D) and Signal strength (S) observables, as well as Magellan specific warnings/indicators
- Generate L1&L2, open for L5 and Galileo bands
- Supporting multiple signal generation for each GNSS
- Optional possibility to extend/reduce data size
- Inserting reference position (static or kinematic) directly inside observation message

GPS, GLONASS, and SBAS observables (as well as static or kinematic reference position) can be generated inside single ATOM transmission. This can give the most compact epoch data presentation. Also the availability of reference position 'tightly coupled' with observables gives very convenient possibility to serve RTK mode against moving base. At the same time, different GNSS data (as well as reference position) can be generated inside different transmissions. In latter case, ATOM generator properly sets Multiple Message Bit allowing compiling complete data epoch.

With multiple signals each modern professional receiver supports, it is very important to specify the type of data escaping any ambiguity in interpretation. ATOM tries to follow RINEX-3 signal naming convention which currently is the widest and most matured identifier for existing and incoming GNSS signals.

The general organization of GNSS observation message is as following:

Message header (time tag, GNSS configuration, etc)
    [GPS data]
    [GLONASS data]
    …
    [Reference position]

In turn each GNSS data are presented as (e.g. GPS):

GPS header (GPS data configuration etc)
    GPS Sat1 data
        Signal1 data
        Signal2 data
        …
        SignalK data
    GPS Sat2 data
        …
    GPS SatN data
        …

where N is the number of Satellites, K is the number of signals for each Satellite.

Signal data include as maximum the following observables (default are **bolded**):

- **Fine Pseudo-Range (C)**
- **Fine Carrier Phase (L)**
- **Signal Strength (S)**
- Fine Doppler (D)
- Warnings (W)

Each observable can be enabled [almost] independently of each other thanks to so called rough range concept.

B. ROUGH RANGE CONCEPT

With proper receiver design, basic observables (pseudo-range and carrier phase) always appear as to be controlled by the same receiver clock. As a result, the 'dynamic' of all pseudo-ranges and carrier phases corresponding to the same Satellite is almost the same. Only ionosphere divergence, receiver biases and some other negligible factors can cause the divergence of one observable against another. This fact is used when generating compact observations. Initially it was introduced for Trimble CMR format [4], later it appeared as a primary concept of standardized RTCM-3 observation messages [2]. Being quite an attractive that time, it became some showstopper nowadays. The problem is that some signal (it is L1 pseudo-range) is selected as 'primary' observable, while all the other ('secondary') signals (e.g. L2 pseudo-range, L1&L2 carrier phase) are generated as the difference against this primary signal.

With multiple signals we have now for each GNSS, it seems such a 'primary-secondary' concept is not convenient. It has at least the following disadvantages:

- Invalid L1 pseudo-range (for whatever reason) automatically leads to inability to present all the other data.
- There is no possibility to send L2 data without sending L1 data. Earlier this was not so important but with current and future availability of L2C and L5, such L1 centered scheme can be ineffective (L5 only receiver can be manufactured in future)
- There is no possibility to send carrier phase data without sending pseudo-range. Carrier phase data have primary interest for precise applications, while [well smoothed] pseudo-range data are usually not needed with the same update rate as carrier phase.

Of course, there already exist some actions to mitigate the negative effect of L1 pseudo-range centered scheme. However, all of them are not so effective compared to the rough range concept used in ATOM.

The idea of rough range concept in ATOM is very simple: each full_range observation #i (pseudo-range or carrier phase) for some particular Satellite #j can be presented as:

full_range(i,j)=rough_range(j)+fine_range(j,i)

i.e. rough_range is unique for given Satellite, while fine_range is unique for each particular observable (pseudo-range or carrier phase). Keeping in mind the worst case ionosphere conditions, fine_range is usually kept within +/-300 meters. As in CMR and RTCM-3, it is assumed that initial integer count is properly removed from carrier phase data at generation startup or after not repaired cycle slip.

With this concept, rough_range itself has not exact physical sense; it is rather some technological value which will be used on decoding side to restore full_range. The same concept is used in ATOM to present Doppler observables. There can be different algorithms to generate rough_range, say:

- Some particular pseudo-range (e.g. L1CA)
- The mean value of all available pseudo-ranges
- Computed range

The natural question arises then: well, it is flexible scheme, but it introduces new (compared to RTCM-3) rough_range field, i.e. increase the size of message. The answer here comes from the fact that there is no any need to present rough_range field with good resolution; on contrary it is sufficient to present it with about 300 meter resolution to allow finally:

- Not to introduce extra bits
- Correct restoring full_range data

When generating corrections (COR) instead of measurements (MES), the room for rough_range is occupied by IODE, while the room for fine_range is occupied by correction itself. This gives finally the same presentation format for COR and MES.

From the above one can see that rough/fine range concept allows finally escaping all the disadvantages related with 'primary-secondary' scheme currently applied in CMR and RTCM-3 messages. Particularly, ATOM can generate:

- Carrier phase without pseudo-range and vice versa
- L2 (or L5) without L1 and vice versa

These 2 opportunities allow effective data decimation to save finally the throughput without loss in final performance.

C. OPTIMIZATION OPTIONS

Nominal configuration of ATOM observation messages (MES and COR) is very similar to standardized RTCM-3 Messages 1004&1012. At the same time, ATOM can be customized to generate 'all' the available observables. At the same time, ATOM can be customized to generate only most important observables and to use compact presentation options.

One of simplest ATOM compact presentation option (NoS) is not sending Signal Strength data in which case ATOM is equivalent to RTCM-3 messages 1003&1011. However, this saving is not so noticeable. The revolution throughput optimization can be made using three options:

- DeCiMating all the data compared to L1 carrier phase (DCM)
- FReeZing identification information (FRZ)
- Using Compact Fine Carrier phase (CFC)

For optimal performance, it is recommended to apply these options for static receiver placed under open sky. This is usually the case with RTK reference station where saving throughput (i.e. bandwidth or/and traffic) is the most important task. At the same time, the two first options can be applied for moving receiver as well; however in this case one can expect some performance degradation (higher percentage of unavailable epochs on rover side). It is because moving receiver is usually affected by cycle slips and constellation changes in higher degree (than static open sky receiver) which in combination with possible short term data link outages can lead potentially to more unavailable epochs on rover side.

On decoder side decimated data can be easily restored provided continuous L1 carrier phase tracking. Restoring pseudo-ranges is trivial even for tens seconds decimation, while to restore decimated L2 (or L5) carrier one has to apply second order estimator to eliminate ionosphere divergence.

When all optimization options are applied, then it appears that identification information (Satellite numbers, Signal identifiers etc) eats the room comparable with observables themselves. In static open sky conditions this identification information is usually changed not so fast. This gives a convenient possibility to freeze (i.e. decimate) most of this information. The idea being the simplest is not trivial in implementation, because one must take into account irregular constellation changes as well as short term data link blockage. The careful freezing implementation in ATOM allows escaping any RTK performance degradation against static open sky reference receiver.

ATOM allows presenting basic observables (pseudo-range and carrier phase) in two different forms: full and compact. In some cases compact form allows saving bits noticeably compared to full presentation. For example, full fine carrier phase takes 24 bits (including carrier itself, cumulative loss of continuity indicator and reserved bits). At the same time compact fine carrier phase takes only 8 bits; this gives 3 times economy without losing final performance. However it should be noted that compact fine carrier phase presentation can be used only for static open sky receiver with between epoch intervals less than about 5 seconds (which is usually the case, because typical reference data rate is 1 Hz).

Each decimation or freezing options allow reducing mean throughput, but at the same time they do not allow reducing peak throughput. At the same time, applying them together but at different epochs, peak throughput can be also reduced. On contrary, compact fine carrier phase allows reducing both mean and peak throughput in equal degree.

We performed a lot of validation of these three compact options between two static open sky receivers simulating RTK function with default ATOM presentation as well as with different optimization options. We made sure that final RTK performance is statistically the same providing ideal data link.

When performing RTK function against static open sky base receiver, one can reach noticeable throughput saving using optimized ATOM observables instead of standardized messages.

It must be emphasized once more that decimation (DCM) and freezing (FRZ) options are implemented in 'adaptive' way, i.e. do not use fixed decimation/freezing intervals, but apply some flexible strategy depending on current situation on reference site. In turn, decoder (on rover side) does not make any a priori assumptions regarding data generation scenario on reference side; on contrary all the information about data presentation form is extracted from ATOM message itself.

THROUGHPUT

Here we give one example (Fig.1) of throughput comparison between standardized RTCM-3 messages and ATOM applying different optimization options. Reference position information is not included into throughput analysis.

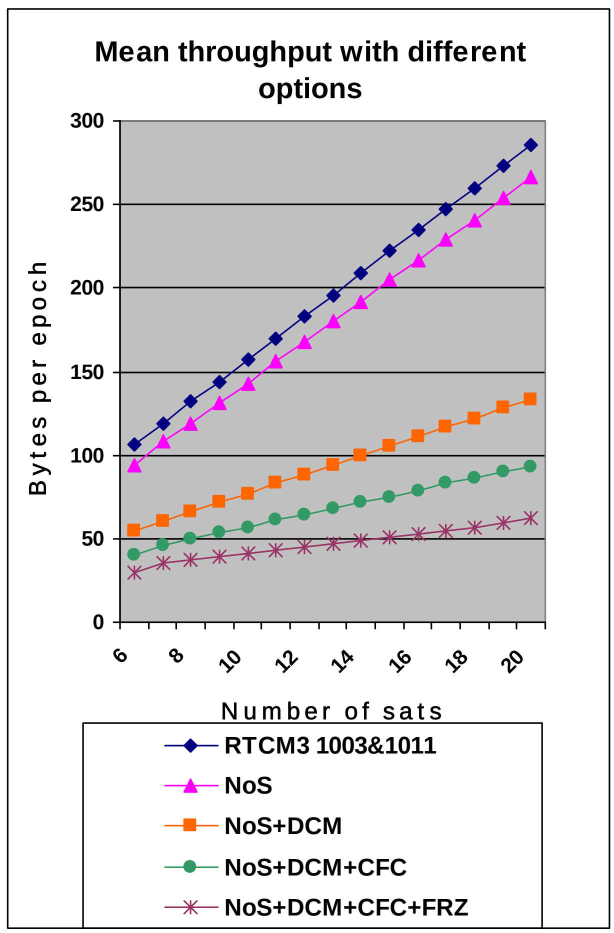

Figure1. Mean ATOM throughput with different options against RTCM3

As a basic variant we selected standardized RTCM-3 message 1003&1011. When we are speaking about the number of Satellites we keep in mind GPS+GLONASS without exact proportion between them (at least one GLONASS is available). This does not affect final numbers because the size of Satellite dependent part is almost the same for GPS and GLONASS for RTCM-3, while it is exactly the same for ATOM.

The mean throughput is measured in bytes per one epoch w/o clear specifying what epoch interval is. However when applying ATOM optimization options, we keep in mind 1Hz update rate which is typical for most reference stations/networks.

One can make sure that ATOM standard compact option (noS) gives approximately the same throughput as 1003&1011. When applying decimation option (DCM) with 5 second interval we get about 2 times economy compared to 1003&1011. And enabling compact fine carrier presentation (CFC) in couple with freezing identification information (FRZ) with 5 second interval we save even more. It is seen that applying full spectrum of ATOM optimization one can reduce mean throughput down to 50 bytes/second (compared to >200 with standardized messages) with 16 Satellites which is typical number for GPS+GLONASS. And primary data update rate is still 1Hz, i.e. this throughput economy is achieved not by increasing epoch interval.

CONCLUSION

Here we presented the overview of new Magellan proprietary GNSS data protocol ATOM. The protocol can be used as the only GNSS data source for different applications. At the same time it can be used in conjunction with existing Magellan proprietary and standardized data protocols.

The usage of standardized transport layer allows not inventing new software to decode ATOM. The optional possibility to encapsulate ATOM allows ease possibility to generate/decode ATOM in single stream together with any other (proprietary and standardized) data.

ATOM observables/corrections appear to be well in line with existing standardized formats like RTCM-3 and RINEX-3. At the same time, ATOM has the following advantages:

- Natural ability to support Multiple GNSS signal data
- Equally well to be used for raw data downloading and differential corrections generation
- Similar presentation form for all GNSS observables
- Escaping any signal-centered strategy
- Natural ability to support exotic applications like RTK against moving base
- Ability to extend/reduce the size of ATOM compared to nominal presentation form

Optimized ATOM messages allow reducing the throughput considerably compared to nominal ATOM configuration and existing standardized protocols. In some particular cases (such as reference static receiver placed under open sky) proper ATOM configuration can give very noticeable saving. This can allow RTK function between two Magellan receivers using either very low band data links or very high update rates. Using GPRS (where user pays for actual traffic) ATOM can give Magellan users noticeable saving with the same RTK performance level.

Here we must address once more data link quality problem when using super compact ATOM options. It is clear that each of three options considered here brings some dependence between ATOM epochs. Once some valuable epoch is missed, few next epochs can in some cases become unavailable. It is proved that with reasonably stable data link (low percentage of missed epochs) ATOM optimization options do not lead to RTK performance degradation. If data link quality became worse (the percentage of missed epochs increases), then one can expect extra percentage of unavailable solutions because of time dependence between epochs. That is why one can reasonably argue about tradeoff between final RTK performance and data link throughput in case of unstable data link. To address this valid concern, we must add the following:

- Often the less data one sends the more reliable communication one has.
- The generation strategy of optimized ATOM is quite ingenious (i.e. do not use simplest fixed interval decimation and freezing) and should not allow too much extra epochs missing even with bad communication.
- ATOM decoder is quite robust and does not make any assumptions about generation strategy and data link quality
- ATOM data restoring strategy is 100% guarded in order not to allow using wrongly restored epoch in RTK computation.
- The communications links we worked with so far did not demonstrate noticeable problems
- The quality and reliability of communication link will be better with time.
- It is the choice for final user either to use compact options (and save throughput in the most of the cases) or use standard ATOM presentation (not saving throughput but be guaranteed that nothing is additionally missed).

ACKNOWLEDGMENTS

Authors would like to thank their Magellan colleagues from Nantes and Moscow for their valuable help with ATOM data testing.

REFERENCES

[1] RINEX. The Receiver Independent Exchange Format, version 3.00, October 3, 2006

[2] RTCM STANDARD FOR DIFFERENTIAL GNSS SERVICES – VERSION 3, RTCM SPECIAL COMMITTEE NO. 104, AUGUST 11, 2006

[3] RTCM STANDARD FOR DIFFERENTIAL GNSS SERVICE - VERSION 2.3, RTCM SPECIAL COMMITTEE NO. 104, AUGUST 20, 2001

[4] Compact Measurement Record (CMR) format, Talbot, 1996