Automated Line-Calling Systems: ITF Evaluation

Revision 21: February 2010

ITF Technical Centre

Supported by

TABLE OF CONTENTS

ITF TECHNICAL CENTRE

AUTOMATED LINE-CALLING SYSTEMS – ITF EVALUATION

1. Introduction

Traditionally, tennis is adjudicated by humans. The decisions as to whether the ball bounced 'in' or 'out' have been made by Line Umpires and Chair Umpires. However, the human information-processing system has limited capacity and there are often occasions when there is uncertainty or dispute over a close call. This is largely because the impact between a ball and surface (a) lasts for a short time (less than 0.01 seconds), and (b) occurs at high speed, with rapid changes in ball position, direction and speed.

Advances in technology have resulted in the development of several methods of automatically determining whether a ball is in or out. The International Tennis Federation (ITF), as the world governing body of tennis, is the focal point for the collation, evaluation and dissemination of information relating to such systems ("automated line-calling systems[1]") to relevant parties.

As major stakeholders in tennis, the ITF, ATP World Tour and WTA Tour are collaborating on the evaluation of automated line-calling systems. These groups recognise that automated line-calling systems offer the potential for more accurate and consistent decision-making than humans, but that their suitability for this purpose must be established before they can be considered for use in professional tennis.

The aim of this document is to describe a process by which the accuracy, precision, reliability, practicality and suitability of automated line-calling systems for use in officiating tennis matches can be established.

2. Theory of ball/surface impact

The impact between a tennis ball and a court surface is complex. The contact area between a ball resting on a court surface is relatively small. If a ball is dropped vertically, the contact area tends to increase as the bottom of the ball deforms (i.e. it flattens out). In general, the greater its velocity in the vertical direction when it hits the surface (up to a point), the more it deforms and the bigger the maximum contact area (see figure 1 for an extreme example). Thus, whether a ball touches a line may be determined by the speed at which it hits the ground, as a faster-moving ball has a greater area of contact with the court surface. For the vast majority of impacts in tennis, the ball is moving horizontally as well as vertically, which increases the complexity of the determination of the true impact location (particularly for human line umpires). Due to the horizontal motion, the ball slides and/or rolls along the ground, and thus leaves the surface at a different position from that at which it first made contact. The amount of sliding changes from shot to shot, and depends on approach

[1] The term 'automated line-calling system' refers to a generic class of tennis equipment designed to determine the location of impact between a ball and court surface, normally with respect to service lines and court boundaries.

angle, speed, spin, and the court surface type, making it difficult to accurately identify the true initial impact point.

Figure 1. The compression of a tennis ball during impact.

Humans may use 'clues' to the impact location. For example, the ball disturbs the surface particles on a clay court, leaving a mark on the surface, while on grass, the presence or absence of chalk dust is sometimes used as evidence of the bounce location. Marks can also be seen on hard courts. In all cases, however, such evidence may be misleading. For example, it is possible that clay particles not directly touched by the ball may be disturbed by those particles moving due to the ball impact, which will result in a larger mark than the true contact area. On a hard court, however, the threshold pressure required between the ball and surface to leave a mark may not be generated until sometime after initial contact, thus leaving a smaller mark than the true contact area. In any case, such information is not available to automated line-calling systems, which (in theory) should produce a more objective decision[2].

3. General system principles

To be acceptable, systems must be:

- Accurate. The system must identify the true impact location of the ball within a specified tolerance.
- Precise. The system must produce the same results for repeated measurements under the same impact conditions.
- Reliable. The system must continue to function without fault or breakdown for extended periods
- Practical. Systems must be easy to use by Chair Umpires and not interfere with other aspects of their role.
- Suitable. Systems must operate over the full range of conditions experienced in tennis, while not exposing players to health and safety risks.

System requirements are described in more detail in Appendix 3. The following section ("*Evaluation Process*") describes the process by which the extent to which a system meets these requirements is evaluated.

[2] When evaluating automated line-calling systems, high-speed video cameras are used as the sole and definitive method by which accuracy is established. The 'mark' left on the court surface following impact is used **only** to aid in the selection of impacts to analyse.

4. Evaluation process

A fundamental principle of the evaluation process is that it is manufacturer-driven. Thus, manufacturers are expected to lead the process in terms of submitting an application and arranging testing. The application forms can be found in Appendices 1 and 2 of this document. The process is summarised as follows:

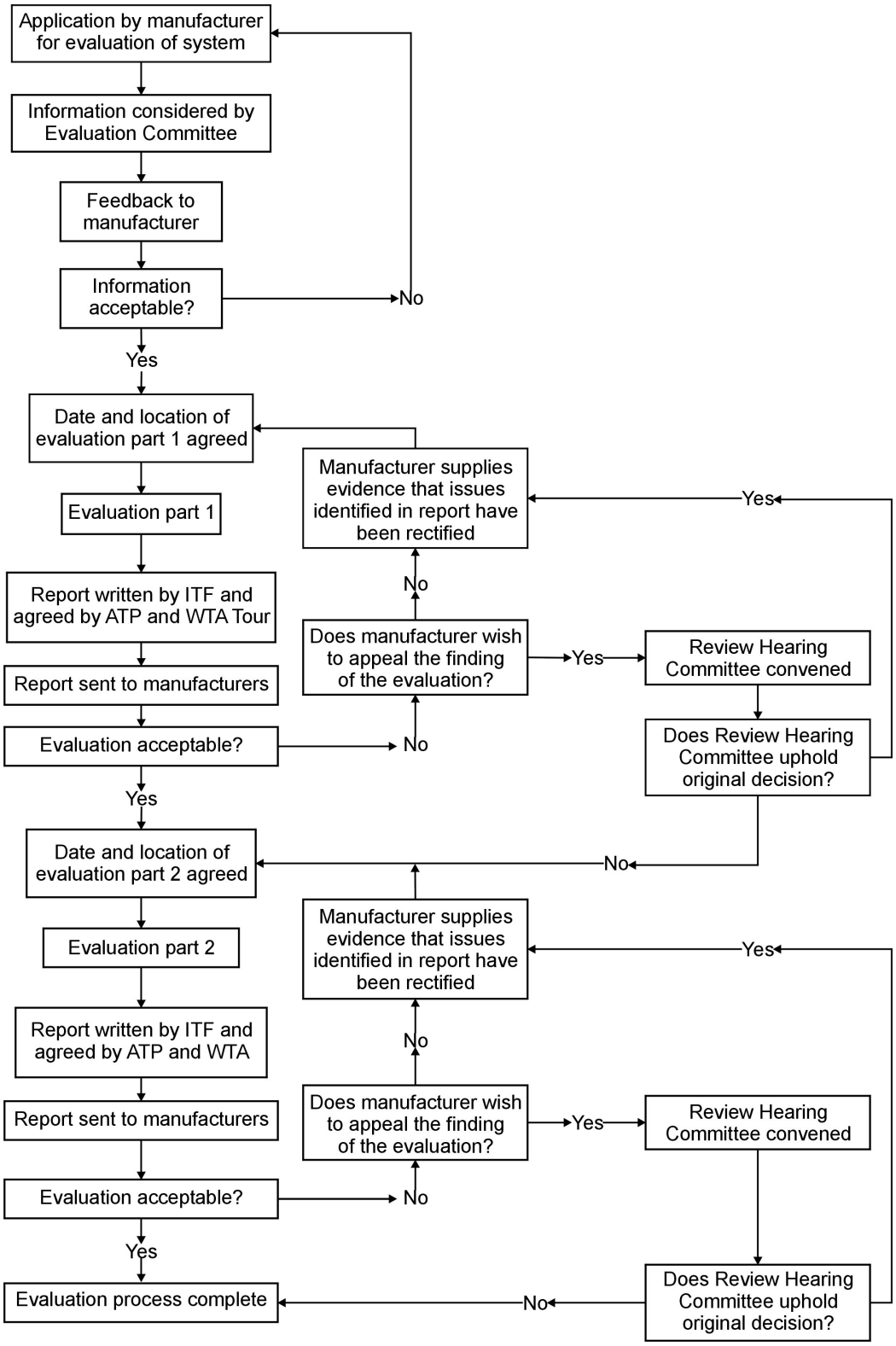


Figure 2. Flow diagram of evaluation process.

i. The manufacturer applies for evaluation of the system (using the forms provided in Appendices 1 and 2). Four copies of a CD/DVD[3] shall also be supplied, which shall contain video of the system (in 'avi' format), showing all equipment, preferably in operation, and all visual devices (e.g. software output, and all information sent to Chair Umpire). An audio commentary is preferable.

ii. Application materials will be circulated to members of the Evaluation Committee for review. Normally, feedback as to the suitability of the application for evaluation will be provided within 4 weeks.

iii. Feedback will be provided if the application is deemed unsuitable for further evaluation. If the application is deemed to be suitable, a time and location for part 1 of the evaluation will be agreed (for which three days should be budgeted). This will not normally be less than 6 weeks from the date of receipt of the application. A backup date should be suggested in the case that the original date is (or becomes) unsuitable.

iv. Evaluation will take place following the protocol for part 1 (see section 4.1). The manufacturer must provide a minimum of 2 players for the duration of the evaluation[4].

v. Following the evaluation, a report agreed by the ITF, ATP World Tour and WTA Tour is compiled. This will normally be sent to manufacturers within 6 working weeks of the test date.

vi. If part 1 of the evaluation is considered to be satisfactory, the manufacturer applies for part 2 of the evaluation[5], for which 5-7 days should be budgeted[6]. If a system fails part 1, the manufacturer may apply for another part 1 evaluation, subject to the Evaluation Committee being satisfied with the responses to the issues identified in the report[7]. In either case, the evaluation will not normally take place less than 6 weeks from the date of receipt of the revised application. A system will not normally be allowed to progress to part 2 of the evaluation process until part 1 has been completed to the satisfaction of the Evaluation Committee.

---

[3] CDs/DVDs are required for a first application, and subsequent applications for which any aspect of the system has been modified, and which can be demonstrated visually.

[4] Players should be of a sufficient standard to aid the evaluation process (such as to hit the ball on, or close to, the lines on request), and must be on hand to play as and when required by the Evaluation Committee. Players will be used at the Evaluation Committee's discretion, but will normally be used only for demonstration of the system operation rather than accuracy testing.

[5] Part 2 requires the system to be installed at a suitable tournament venue, for which the manufacturer is responsible. See Appendix 3. Assistance with contacts and access to tournaments may be provided by the Evaluation Committee.

[6] The number of days depends on the nature of the evaluation. Five days should be budgeted for a 'shadow' test. In exceptional circumstances, the Evaluation Committee will allow a combined 'shadow' and 'live' test, for which 7 days should be budgeted. [An **additional** 7 days should be budgeted for the 'live' tests when it is separate from the 'shadow' test]

[7] The Evaluation Committee requires evidence of the action taken following a failed evaluation before another evaluation will be scheduled. This must include a written description of the problem, how its cause was identified, how it was rectified and evidence that the correction action was successful.

vii. A test date and location is agreed.

viii. Evaluation takes place in accordance with the protocol for part 2 specified in section 4.1.

ix. Following the evaluation, a report is agreed by the ITF, ATP World Tour and WTA Tour. This will normally be sent to the manufacturer within 6 working weeks of the test date.

x. A system is approved when it passes all of part 2 of the evaluation. If a system fails part 2, the manufacturer may apply for a further part 2 evaluation[8]. This will not normally be less than 6 weeks after the previous evaluation, subject to the Evaluation Committee being satisfied with the responses to the issues identified in the report from the failed part 2 evaluation[9].

4.1 Evaluation protocol

The evaluation itself is divided as follows:

4.1.1 Part 1 – accuracy and familiarisation

Stage 1 (accuracy) – normally 1-2 days

A key element of the evaluation is the extent to which the system is able to identify the true ball impact location. This will be evaluated based on simulated impacts that are in the range possible during normal play and will normally include various impact locations and conditions. System performance will be assessed in terms of (as applicable[10]):

- Reliability of tracking/detection – the percentage of times that the system successfully tracks/detects the ball.
- Decision-making success rate – the percentage of correct in/out calls (within tolerance)[11].
- Average accuracy per line – the average of the absolute discrepancies for an individual line.
- Minimum accuracy – the limiting value of accuracy within which all 'in'/'out' calls are made correctly.
- Single impact accuracy – the maximum single discrepancy for all impacts.

[8] It should be noted that the Evaluation Committee will take the overall performance of a system into account when deciding whether a system passes each part of the evaluation.

[9] The Evaluation Committee requires evidence of the action taken following a failed evaluation before another evaluation will be scheduled. This must include a written description of the problem, how its cause was identified, how it was rectified and evidence that the correction action was successful.

[10] Accuracy measures used will depend on system type.

[11] It is not always possible to produce impacts at the ideal locations, and so aspects of decision-making and accuracy may be inferred from interpolation of actual data.

Stage 2 (familiarisation) – normally 1 day

- The system is installed on one court at a selected venue.
- The manufacturer conducts a tour of the system characteristics and performance.
- The manufacturer demonstrates the system during operation during play (normally 1 × 3-set match[12]).

N.B. Accuracy testing should take place outdoors where possible. For an indoor test, the minimum vertical illuminance required at court level is 2000 Lux.

4.1.2 Part 2[13]

4.1.2.1 Accuracy and 'shadow' test

This part of the evaluation process normally takes place at a tournament (on 1 court only) and will only be undertaken once part 1 has been completed to the satisfaction of the Evaluation Committee, and when evidence that any recommendations from part 1 have been implemented. The primary aims are to perform the following:

- Accuracy test – repeat part 1 (stage 1) to demonstrate day-to-day accuracy[14].
- Shadow test – the system is used in parallel with the on-court officials.

A total of 5 days should be budgeted for this part of the process, to include 1-2 days of accuracy testing, 2-3 days of shadow testing, and 1 fall-back day. It is not essential for the shadow test to take place at an organised event, although enough players and officials must be provided to allow 2-3 days of continuous play[15].

4.1.2.2 Accuracy and 'live' test

If the system performance in accuracy and shadow testing is deemed suitable, evaluation proceeds to a 'live' test, in which the system is used by officials to make/review on-court decisions[16]. The system must be installed at a professional event (e.g. ITF Pro Circuit). This aspect of the evaluation will focus on:

- Practicality of the system for use by Chair Umpire[17].
- Reliability and suitability of the system.

---

[12] The duration of play may vary according to the needs of the Evaluation Committee.

[13] The Evaluation Committee reserves the right to conduct accuracy testing at any stage of a shadow and/or live test.

[14] Approval will require a system to be evaluated for accuracy under natural, artificial and changing illumination.

[15] Players must be of similar quality to top-ranked players.

[16] This may or may not be at the same event as for the shadow test [ideally different].

[17] As applicable; the requirements for review and real-time systems will be different. See Appendix 4.

A total of 7 days should be budgeted for this part of the process, to include 1-2 days of accuracy testing, and 3-4 days of live testing.

4.1.2.3 Accuracy and Shadow/Live test

Under exceptional circumstances[18], the Evaluation Committee will allow the shadow test (A3.2.1) and live test (A3.2.2) to be conducted in immediate succession. **N.B. all such combined shadow/live tests must take place at a professional tennis event to be agreed by the Evaluation Committee**. A total of 7 days should be budgeted for this part of the process, which will normally include 1-2 day of accuracy testing, 2 days of shadow testing, and 3-4 days of live testing.

IMPORTANT NOTES

All costs associated with evaluation of systems are to be met by manufacturers. This includes (without limitation): travel, accommodation and subsistence of all Evaluation Committee members; transport and hire of equipment (as necessary), and; time (e.g. for report writing). All applications for evaluation must be accompanied by written acceptance of the costs of evaluation.

A system will not be deemed to have completed the evaluation process until it has performed to the Evaluation Committee's satisfaction in all parts of that process.

The Evaluation Committee reserves the right to consider the performance over all evaluations in its deliberations on whether that system is deemed to have completed the evaluation process.

Neither participation in, or completion of, the evaluation process is any guarantee that a system will be adopted in tournament play.

This procedure is subject to ongoing review and may be amended as and when necessary and without prior notice.

NOTE TO END USERS: Completion of the process implies that a system has attained the level of performance specified in the relevant report(s) on the day(s) of evaluation, and at the location(s) (including facility and surface) and under the prevailing ambient conditions on those days. It is strongly recommended, therefore, that end users establish that the system performs to the required standard in the location (and, if possible, under the ambient conditions) in which the end user intends it to be used. The ITF, ATP World Tour and WTA Tour will under no circumstances be responsible in the event of any failure or poor performance of a system.

[18] For a combined shadow and live test to be considered, the system will normally either have been free of faults in all previous parts of the evaluation process, or have met the required standards in both accuracy and shadow testing (but not necessarily during a single evaluation) on a previous occasion.

APPENDIX 1 – APPLICATION FOR EVALUATION OF AN AUTOMATED LINE CALLING SYSTEM

Contact Information

Product name:	
Applicant company:	
Contact person(s):	
Address:	
Telephone:	
Mobile phone:	
Email:	

Application Checklist

Completed forms ITF/ALCS/01 & ITF/ALCS/02	
4 CDs/DVDs	
Annotated visual evidence of equipment locations ITF/ALCS/03	
Written confirmation of acceptance of responsibility for the costs of evaluation	

Declaration

I have read and understood the contents of this document "Automated Line-Calling Systems: ITF Evaluation" (revision 21). Signed on behalf of applicant company:		
Name (please print):		Electronic Signature:
Date:		

Completed application to be submitted to:

ITF Technical Centre, ITF Licensing UK Ltd,
Bank Lane, Roehampton, London, SW15 5XZ, UK
Tel: +44 (0) 208 878 6464
Fax: +44 (0) 208 392 4773
Email: stuart.miller@itftennis.com

APPENDIX 2 – SYSTEM INFORMATION[19]

A2.1 General description

System type ('real-time', i.e. designed to make line calls immediately, or 'review', i.e. designed to provide a replay as an aid for the Chair Umpire)[20].
System coverage (whole court or service-line only).
Is system claimed to work on all surfaces (if 'no', please specify surfaces for which system is not designed).
Impact detection method (e.g. light beam, pixel recognition, court marking).
Description of the major system components.
Cost (to buy outright and to rent) and technical support provided.
Portability (dimensions, weight).
Claimed advantages of system (compared to other systems).
How long does the system take to set up (please divide into installation time and calibration/ verification, and state whether these can be done in parallel)?

[19] If you have previously made an application, please highlight the items that have changed since the last application, by adding the text "[**change**]" at the end of the relevant section(s)

[20] If a system offers the capability of functioning as both a 'real-time' and 'review' system, and the applicant wishes the system to be evaluated for both functions, this shall be indicated here.

Minimum/preferred number of people required to operate system.
Number of courts on which the system has been installed, and period(s) of installation.
Has the system been used by players?

A2.2 Hardware

Type (manufacturer and model), number (minimum and optimum) and size of detection devices.
Local requirements of detection device(s).
Number of PCs required and location(s) with respect to detection devices, court surface and operators.
Installation requirements (including power requirements, any modifications to court, connections between components).
Are detection devices or hardware exposed to interference by spectators, officials, etc.?
Is a backup system available in case the main system fails (e.g. through power loss or hardware malfunction)?
Number of access points to hardware during normal use (i.e. by spectators, officials, etc.).

Number of failure points (i.e. hardware locations where the normal system function may be disrupted).
Is any part of the system laid under the court surface (please specify)? If so, please state how many times a court can be resurfaced before system performance (e.g. accuracy) is affected.

A2.3 Software

General description.
If camera-based, frame rate and shutter speed of camera.
If camera-based, please describe software algorithm used for reconstruction of ball trajectory.
If camera-based, how many images of a trajectory are needed to maximise tracking accuracy (specify pre- and post-impact separately)?
Does system use post-bounce data to establish ball/surface contact locations?
Information provided to Chair Umpire (including time delay between ball contact and signal to Chair Umpire).
(How) does system deal with obstruction of ball by player(s)?

(How) does system deal with the non-circularity of ball-surface contact area?
(How) does system deal with sliding/rolling of ball during impact?
(How) does system transfer data into other electronic systems (e.g. statistical databases and web-based 'live' scoring transmissions)?

A2.4 Performance

Please note that any performance characteristics supplied below will only be accepted in lieu of testing in exceptional circumstances. Where performance is not known, specify 'not known'.
Accuracy (average difference between true contact point and that recorded). Specify number of test impacts and inbound ball conditions.
Can system accuracy be displayed by the software and how is it determined?
Resolution (i.e. smallest determinable change in ball position).
Precision (for a constant input, system makes the same decision).
What is the longest period the system has operated uninterrupted by faults?
Range of temperature and humidity within the system operates at maximum accuracy.
Does the system produce the same accuracy under natural light and floodlights? What is the minimum lighting (in lux) the system requires to operate accurately?

Does the system function with the same accuracy at all venues?
Can the system perform a self-diagnosis of accuracy during use?
Has performance specified above been established by manufacturer or independently (if independently, provide contact details of testing body, and supply a copy of any test reports)?

A2.5 Safety and security

Does the system generate any emissions that are above ambient levels (such as electrical or magnetic fields)? If so, please specify type and intensity.
Does the software contain security features to (a) prevent hacking, and (b) ensure that data are not manipulated, either accidentally or deliberately? If so, please describe.

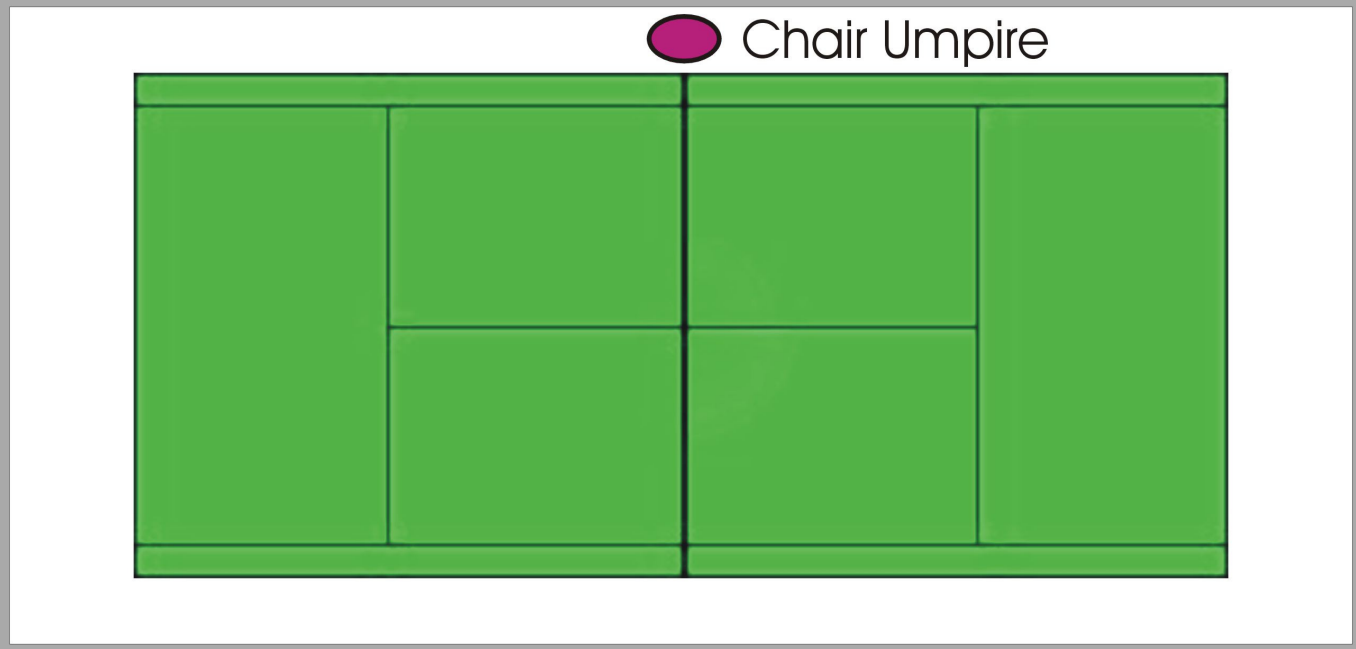
Please indicate the preferred location of all equipment (including connecting wires) by drawing on, and labelling, the diagram at the end of this section.

A2.6 Other information

Please list any collaborations and/or partnerships with companies other than that names in Appendix 1. Please specify the aspects of the system for which the companies specified are responsible.
Please provide any further information that you feel is of relevance here.

Note: Information provided in Appendix 2 may be used in the evaluation report. Photographic information/evidence for any aspect of system information is welcomed.

Seating Area



APPENDIX 3 – SYSTEM CHARACTERISTICS TO BE EVALUATED

To satisfy the principles described in section 3, systems need to have the characteristics specified in sections A3.1-A3.7 below. **The extent to which these characteristics can be evaluated in this process may vary due to (e.g.) the characteristics of the system, and time and geographical constraints.**

A3.1 General

The line-calling system must provide accurate and reliable line calls in a timely manner.

The system must provide easy manual override, whenever so required.

The system must provide full performance under any combination of conditions that can be expected in a tennis match.

The line-calling system must be unobtrusive when installed in its operating environment. At a minimum, it must not adversely affect the players or the officials.

The system must not occupy so much space courtside as to interfere with the court layout, infrastructure or operation in comparison with currently accepted practice. It must not draw undue or excessive attention and must not detract from the enjoyment of spectators.

The system must not adversely affect activities related to television or other communication media.

The system should be controllable by the Chair Umpire, whether directly or indirectly.

Information must be provided in a timely manner, such that no interference is caused with the smooth running of the match.

A3.2 Performance

The system should provide complete coverage of all court lines.

The system must be operational under natural-, artificial- and variable-light conditions that are suitable for play.

The system performance must be unaffected by rain or moisture on the court surface, or surface temperatures in the range of 0-70º Celsius.

The system must be operational under all weather conditions that are deemed suitable for play by the Referee.

The system must meet the required performance standards under all combinations of normal playing conditions, including variation in bounce location, ball speed and trajectory angle.

The system will not be affected by the presence or movements of players (singles and doubles), including stepping or standing on the lines.

The system performance will not be affected by any objects worn or carried by players or officials, including rackets.

A review system should provide a visual image of the ball trajectory and/or footprint relative to the lines on the court. The image quality and accuracy will be sufficient to visually confirm the system's (automated) decision. This image will be available to the Chair Umpire and, upon approval by the ITF, to others such as spectators or television viewers.

A real-time system should provide "out" calls automatically and simultaneously to umpires, players and spectators in a clear and unambiguous way with an audible signal. A number of choices of types and volumes of audible signals (including synthetic voice) must be available for use by the Chair Umpire or tournament management. The time interval between ball bounce and the audible signal must not be significantly longer than that normally required for a Line Umpire's call.

A3.3 Court and Balls

If the court must be modified in any way, the modified court must not be materially different from a standard court in appearance and/or all playing characteristics.

The playing characteristics of the court must be the same whether or not the line-calling system is in operation.

If the ball must be modified in any way, all such modifications must be applicable to all ball types in current use. Modified balls must be ITF Approved.

The sensing characteristics of the system/ball combination must not deteriorate over the ambient temperature range of 0-50º Celsius, or by the appearance of moisture in the cloth cover.

There must be a sufficiently wide 'sensed zone' either side of the line. The minimum such zone is from 20 cm 'in' to 60 cm 'out'.

A3.4 Accuracy and Reliability[21]

The decision-making success rate (i.e. "in" or "out" decisions) for all impacts in the range 100 mm inside the line to 50 mm outside the line (as measured by the ITF) should be 100% with a tolerance of ± 5 mm.

The average absolute discrepancy between the system and the ITF for all impacts on a single line on court should not exceed 5 mm.

The maximum discrepancy between the system and the ITF for all impacts should not exceed 10 mm.

[21] The specifications described in this section apply to balls that legally cross a line from outside to inside.

A3.5 Control/Replay Equipment

The line-calling system must include equipment to enable a Chair Umpire to control and/or monitor (or be aware of) its status, and that of the match.

The control functions of any handset used by the Chair Umpire must be compatible with existing scoring handset systems used in professional tennis and must have suitable interface characteristics.

When used as a replacement for on-court Line Umpires, the system status monitoring must be provided to the Chair Umpire in an unobtrusive manner, without affecting the task of officiating.

When used as a replacement for on-court Line Umpires, the control equipment must have the capability to allow the use of special purpose software (e.g. for visual confirmation of all decisions or actions taken).

A3.6 Reliability and Maintenance

The mean time between any breakdown should be not less than 80 hours of operation.

All parts of the system should be modularised to facilitate failure repair.

The system should have multiple (at least double) redundancies built into its design.

The system should constantly and automatically self-test. As a minimum, it should have a built-in capability for automatic self-test, which should be carried out at frequent intervals during standby conditions.

Following initial (pre-event) calibration, the system should be capable of re-calibration at that event during the course of normal play (i.e. between successive matches) without people or objects on the court surface.

The system should have embedded test diagnostics.

The majority of failures of the system should be easily correctable by the Chair Umpire or a trained technician within 5 minutes.

For systems using impact-sensing elements under the court surface, a failure in one of those line elements must not affect the performance of other elements or any other part of the system.

The system must have a high level of discrimination capability against non-ball signals. This would include background rejection, and interference effects due to magnetic, optical, infrared or other sources.

There must be an adequate level of protection against deliberate or inadvertent actions by spectators to either produce a false or incorrect signal (such as an audible “out” beep) or otherwise interfere with the system.

A3.7 Safety and Interference

The system must not present a risk to players due to collision with equipment, electric shock, etc.

Systems using energy sources such as electromagnetic radiation, infrared beams or

optical methods should not pose any risk to the players, officials or spectators.

The system must not have any significant adverse effects on other nearby electrical equipment, such as TV cameras, data transmission lines or microphones.

Systems must be capable of operating satisfactorily in the electromagnetic environments normally found at major tennis venues.